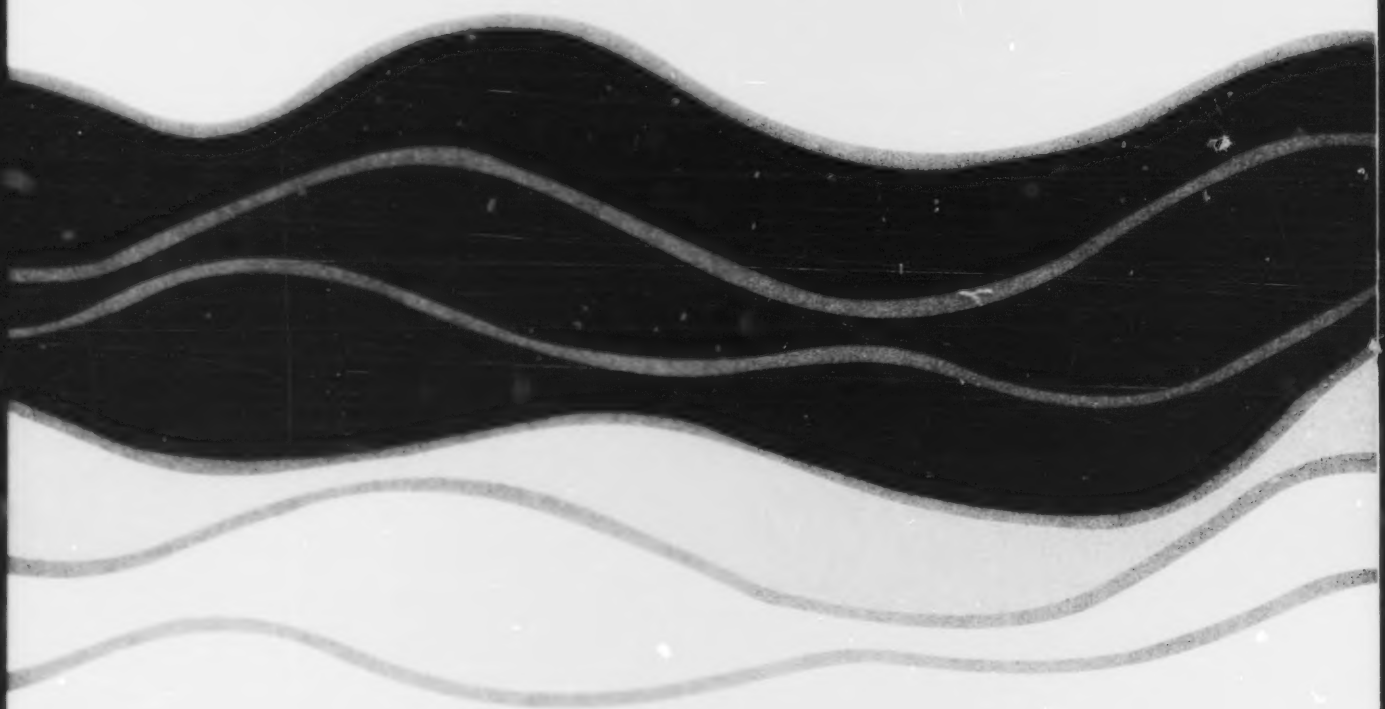
NATIONAL WATER RESEARCH INSTITUTE

INSTITUT NATIONAL de RECHERCHE sur les EAUX



QUALITY ASSURANCE GUIDELINES FOR BIOLOGY IN AQUATIC ENVIRONMENT PROTECTION

QUALITY ASSURANCE GUIDELINES FOR BIOLOGY IN AQUATIC ENVIRONMENT PROTECTION

Environment Canada
Conservation and Protection
National Water Research Institute
Burlington, Ontario
Canada L7R 4A6

Report Prepared by Beak Consultants Ltd.
14 Abacus Road, Brampton, Ontario
Canada L6T 5B7

92-01716/

ACKNOWLEDGEMENTS

This Quality Assurance Guidelines for Biology in Aquatic Environmental Protection document was prepared from input received from members of the Organizing Committee and participants at the Workshop. Their contribution and effort are gratefully acknowledged.

Dr. D. Hart, Beak Consultants Limited, who was contracted to facilitate the Workshop and prepare the final report deserves our special thanks.

Copies of this report are available from:

National Water Research Institute
867 Lakeshore Rd., P.O. Box 5050
Burlington, Ontario L7R 4A6
ATTN: Alfred S.Y. Chau

MANAGEMENT PERSPECTIVE

The implementation of legislation such as the Canadian Environment Protection Act, CEPA, and the Ontario Municipal Industrial Strategy for Abatement, MISA, has increased the need to use biological criteria to assess the impact of contaminants and effluents on the environment. These Quality Assurance Guidelines for Biology in Aquatic Environmental Protection were prepared in response to a need identified at the Alliston Workshop on Biology in the New Regulatory Framework for Aquatic Protection, April 1988 organized by Environment Canada.

An Organizing Committee of a small group of experts met at the National Water Research Institute, September 28, 1988, to develop a first draft of the guidelines. A two day workshop was held February 20-21, 1990 to critically review the revised document. Dr. D. Hart, Beak Consultants Limited was contracted to prepare the various drafts of the guideline and to serve as coordinator of the Workshop.

This current version of Quality Assurance Guidelines for Biology in Aquatic Environmental Protection has benefitted significantly from the deliberations of the workshop and from the input of numerous knowledgeable reviews. Due to the diversity of the subject, the document is intended to serve only as a generic guide and to be expanded by users to meet specific objectives and requirements. It is also intended that future refinements and improvements will be made as the state of knowledge of biological quality assurance is developed.

John Lawrence
Director
Research and Applications Branch

April, 1991

PERSPECTIVE-GESTION

La mise en application d'une législation comme la Loi canadienne sur la protection de l'environnement et la Stratégie municipale et industrielle de dépollution (SMID) de l'Ontario, a rendu plus pressante la nécessité d'utiliser des critères biologiques pour l'évaluation de l'impact des contaminants et des effluents sur l'environnement. Les présentes Lignes directrices sur l'assurance de la qualité visant la détermination de paramètres biologiques dans le domaine de la protection de l'environnement aquatique ont été préparées en réponse à un besoin qui a été reconnu au cours de l'Atelier intitulé *Workshop on Biology in the New Regulatory Framework for Aquatic Protection* qui a eu lieu à Alliston en avril 1988 et qui a été parrainé par Environnement Canada.

Un Comité organisateur formé d'un petit groupe d'experts s'est réuni le 28 septembre 1988 à l'Institut national de recherches sur les eaux pour préparer une première ébauche de lignes directrices. Le document révisé a ensuite fait l'objet d'un examen critique au cours d'un atelier de deux jours qui a eu lieu les 20 et 21 février 1990. Les services de M. D. Hart de Beak Consultants Limited ont été retenus pour la préparation des diverses ébauches de lignes directrices ainsi que pour la coordination de l'atelier.

La version actuelle des Lignes directrices sur l'assurance de la qualité visant la détermination de paramètres biologiques dans le domaine de la protection de l'environnement aquatique a été grandement enrichie par les discussions qui se sont déroulées au cours de l'atelier et par les nombreux examens critiques très éclairés. En raison de la grande diversité des sujets abordés, le document doit être considéré comme un guide général que les utilisateurs auront à adapter en fonction de leurs objectifs ou leurs besoins précis. Il est également prévu que le document sera revu et modifié au fur et à mesure que les connaissances dans le domaine de l'assurance de la qualité appliquée à la biologie progresseront.

Le Directeur
John Lawrence
Direction de la recherche pure et appliquée

EXECUTIVE SUMMARY

At the Alliston Workshop on Biology in New Regulatory Framework for Aquatic Protection, April 1988 organized by Environment Canada, it was recognized that there was a need to develop general guidelines for the design and implementation of quality assurance programs for biological measurement systems. The current guidelines were developed in response to this need. They are the effort of biological and QA experts from federal, academic and industrial sectors across Canada.

This document describes the basic elements of an internal QA/QC program based on the principles of good laboratory practice, and then goes on to describe the application of these principles in specific biological subject areas. Topics include aquatic toxicity testing, benthic community assessment, microbiological assessment and paleolimnology. A QA checklist for toxicity testing is provided which draws attention to key areas such as water supply, animal controls, dose selection and LC50 calculation that are known to affect data quality.

The need for initiation of interlaboratory QA studies and follow-up investigations is identified. Method inconsistencies associated with taxonomy, sample seasons and methodologies, preservation and sorting techniques which often limit the suitability of data for temporal and spatial comparisons are also addressed. Sample replication is recommended to permit estimation of total precision, while split sample and reference collection exchanges may be used to quantify interlaboratory comparability. A checklist is provided.

RÉSUMÉ ADMINISTRATIF

À l'Atelier intitulé *Workshop on Biology in the New Regulatory Framework for Aquatic Protection*, qui a eu lieu à Alliston en avril 1988 sous les auspices d'Environnement Canada, on a reconnu la nécessité d'élaborer des lignes directrices générales pour la conception et la mise en oeuvre de programmes d'assurance de la qualité pour les systèmes de mesures biologiques. Les lignes directrices actuelles viennent répondre à ce besoin. Elles sont le fruit des efforts d'experts dans les domaines de la biologie et de l'assurance de la qualité provenant de divers secteurs (gouvernement fédéral, universités, industrie) et de partout au Canada.

Le présent document décrit les éléments fondamentaux d'un programme interne d'AQ-CQ fondé sur les principes de la bonne pratique de laboratoire ainsi que l'application de ces principes dans des domaines précis de la biologie. Parmi les questions abordées figurent les tests de toxicité dans le milieu aquatique, l'évaluation des communautés benthiques, l'évaluation microbiologique et la paléolimnologie. Le document contient une liste de vérification AQ applicable aux tests de toxicité et qui attire l'attention sur des éléments clés qui influent sur la qualité des données, tels l'approvisionnement en eau, les témoins animaux, le choix des doses et le calcul de la CL_{50}.

La nécessité d'entreprendre des études d'AQ interlaboratoires et des études de suivi y est soulignée. Il est également question des variations touchant les méthodes liées à la taxinomie et à l'échantillonnage (moment et méthodes, conservation et tri des échantillons) qui, souvent, limitent l'utilité des données pour les comparaisons temporelles et spatiales. La répétition de l'échantillonnage est recommandée pour permettre l'estimation de la précision globale, et il est indiqué que le recours à l'échantillonnage fractionné et à des échanges de collections de référence peut servir à établir la comparabilité interlaboratoires. Une liste de vérification est fournie.

ORGANIZING COMMITTEE

Mr. Alfred S.Y. Chau (CHAIRMAN)
Chief, Quality Assurance Project
National Water Research Institute
BURLINGTON, Ontario

Dr. Hugh Eisenhauer
Director, Toxic Chemicals Program
Conservation and Protection
HULL, Quebec

Dr. Kristin Day
Contaminants Pesticides Projects
Rivers Research Branch
BURLINGTON, Ontario

Dr. Martin Samoiloff
Bioquest International Inc.
WINNIPEG, Manitoba

Dr. David Rodgers
Biology Research Section
Research Division
Ontario Hydro
TORONTO, Ontario

Dr. Dieb Birkholz
Enviro-Tes Laboratories
EDMONTON, Alberta

Mr. Gary Sergy
A/Head, Technology Development
& Technical Services Branch
Conservation and Protection
Western and Northern Region Office
EDMONTON, Alberta

Dr. Peter Wells
Senior Advisor, Marine Environment
Quality
Conservation and Protection
Atlantic Region
DARTMOUTH, Nova Scotia

Mr. Ken Doe
Head, Aquatic Toxicology Section
Conservation and Protection
Atlantic Region
DARTMOUTH, Nova Scotia

Dr. John Lawrence
Director, Research & Applications
Branch
National Water Research Institute
BURLINGTON, Ontario

CONTRIBUTORS AND WORKSHOP PARTICIPANTS

Dr. D.R. Hart
Beak Consultants Limited
BRAMPTON, Ontario

Dr. Robert Prairie
Centre de Technologie Noranda
POINTE-CLAIRE, Quebec

Dr. Brian Cumming
Queens University
KINGSTON, Ontario

Dr. John Kingston
Queens University
KINGSTON, Ontario

Dr. Jim Pollard
Lockheed Engineering
and Sciences Co.
LAS VEGAS, Nevada U.S.A.

Dr. Dennis Delorme
National Water Research Institute
BURLINGTON, Ontario

Mr. R. Scroggins
Industrial Programs Branch
Environment Canada
OTTAWA, Ontario

Dr. Hugh Eisenhauer
Environment Canada
Management and Emergencies
HULL, Quebec

Dr. Martin Samoiloff
Bioquest International Inc.
WINNIPEG, Manitoba

Dr. Dieb Birkholz
EnviroTest Laboratories
EDMONTON, Alberta

Dr. David Rodgers
Ontario Hydro
Biological Research Section
TORONTO, Ontario

Mr. Mike Walsh
Ontario Ministry of the Environment
TORONTO, Ontario

Mr. Marius Marsh
Ontario Ministry of the Environment
Phytotoxicology Laboratory
BRAMPTON, Ontario

Dr. Ian Davies
Freshwater Institute
WINNIPEG, Manitoba

Dr. Peter Wells
Atlantic Region
Environment Canada
DARTMOUTH, Nova Scotia

Dr. Kristin Day
National Water Research Institute
BURLINGTON, Ontario

Mr. Barney Dutka
National Water Research Institute
BURLINGTON, Ontario

Dr. Shirland Daniels
National Water Research Institute
BURLINGTON, Ontario

TABLE OF CONTENTS

		Page
1.0	INTRODUCTION	1.1
2.0	GOOD LABORATORY PRACTICE	2.1
3.0	ELEMENTS OF A QA/QC PROGRAM	3.1
3.1	Responsibilities and Authorities	3.1
3.2	Data Quality Objectives	3.2
3.3	Sampling Procedures	3.3
3.4	Analytical Procedures	3.5
3.5	Quality Assurance Procedures	3.6
3.6	Documentation Procedures	3.10
4.0	QA IN AQUATIC TOXICITY TESTING	4.1
4.1	Sample Collection and Storage	4.1
4.2	Records of Test Organisms	4.2
4.3	Test Methods	4.3
4.4	Quality Control	4.5
4.5	Interlaboratory Comparison	4.7
4.6	Toxicity Degradation Studies	4.8
4.7	Raw Sample Compliance Tests	4.10
4.8	Multi-species Toxicity Tests	4.11
4.9	*In Situ* Bioassays	4.12
4.10	Laboratory Evaluation Checklist	4.14
4.11	Residue Toxicity and Bioaccumulation	4.15
5.0	MICROBIOLOGICAL QA	5.1
5.1	Sample Collection and Storage	5.1
5.2	Sterilization	5.1
5.3	Methods and Records	5.2
5.4	Quality Control	5.2
5.5	Interlaboratory Comparison	5.3
5.6	Viral Testing	5.3
5.7	Microbial Toxicity Testing	5.4

		Page
6.0	QA IN AQUATIC FIELD STUDIES	6.1
6.1	Preliminary Survey and Design	6.1
6.2	Training and Performance Evaluation	6.2
6.3	Sample Collection, Preservation and Storage	6.3
6.4	Additional Effort	6.3
6.5	Sample Replication	6.4
6.6	Duplicate Processing and Exchange	6.4
6.7	Sorting Verification	6.5
6.8	Taxonomic Verification and Archives	6.5
6.9	Data Handling and Reporting	6.6
6.10	Rapid Bioassessment	6.6
6.11	Macrophyte Surveys	6.7
7.0	QA IN PALEOLIMNOLOGY	7.1
7.1	Responsibilities and Authorities	7.1
7.2	Sampling Procedures	7.2
7.3	Analytical Procedures	7.5
7.4	Quality Control	7.6
7.5	Documentation Procedures	7.10
8.0	REFERENCES	8.1
APPENDIX 1:	Specific Examples	
APPENDIX 2:	Quality Assurance Glossary	

LIST OF TABLES

Table No.

4.1: Quality Assurance Checklist for Toxicity Laboratories

6.1: Quality Assurance Checklist for Aquatic Field Studies

A3.1: Examples of Sampling Designs
A3.2: Estimation of Replicates Required
A3.3: A DQO Statement Format for a Compliance Study
A3.4: A DQO Statement Format for a Spatial Trend Study

A4.1a: Recommended Preparation of Reconstituted Fresh Water
A4.1b: Specifications for Alternative Dilution Water
A4.1c: Recommended Preparation of Reconstituted Sea Water

A5.1: Autoclave Sterilization Temperature and Time
A5.2: Specifications for Membrane Filters and Absorbent Pads
A5.3: Monitoring Laboratory Equipment
A5.4: Maximum Storage Times for Prepared Media
A5.5: Criteria for Laboratory Pure Water Used in Microbiological Procedures
A5.6: Control Cultures for Standard Microbiological Tests

LIST OF FIGURES

Figure No.

3.1: Range and Average Charts for Control of Precision and Accuracy
3.2: Range Chart for Response-Dependent Precision Control
3.3: Criterion and Limit of Detection

4.1: Control Charting of Duplicate Toxicity Test Data
4.2: Average Control Charts for Reference Toxicants

Figure No.

6.1: Designing Effort to a Species Inventory Objective
6.2: Sequential Sampling to Meet a Precision Objective

7.1: Custody Record - Lake Sediment Samples

1.0 INTRODUCTION

In April 1988, the Alliston Workshop was organized by Environment Canada, Conservation and Protection, to address issues related to the use of biological methods for aquatic environmental protection (Day et al., 1988). Recent federal and provincial legislative and policy initiatives are providing an updated regulatory framework for expanding use of biological testing in environmental protection. The workshop concluded that such regulatory applications would require development of QA/QC guidelines for biological testing, and recommended that the Federal government should provide guidance and leadership in this area. Biological testing is broadly interpreted here to include both laboratory toxicity testing and in situ biomonitoring.

This document is intended to provide guidance to both government and private laboratories in the design and implementation of quality assurance programs for biological measurement systems, and to scientific authorities in evaluation of contract laboratories. The fundamental principles of quality assurance are the same in biological and chemical laboratories. However, since some potential sources of error are different, some details of implementation must be unique to the biological measurement system. Moreover, standards suitable for assessment of accuracy are often unavailable for biological measurements.

Quality assurance (QA) encompasses a wide range of both management and technical practices designed to ensure an end product of known quality commensurate with intended use of the product. In scientific studies, the end product is a set of data which, upon analysis and interpretation, will be used to draw conclusions and support decisions. In environmental science, the decisions typically involve either pollution abatement or new regulatory initiatives.

Quality control (QC) is one aspect of quality assurance. It includes the techniques used to measure and assess data quality and the remedial actions to be taken when data quality objectives are not realized. Assurance of adequate data quality is only possible when specific data uses and Data Quality Objectives have been defined. Users of the data must play a lead role in defining Data Quality Objectives for a study and ascertaining whether laboratory quality control limits are consistent with these objectives.

A complete quality assurance program must encompass both sampling and analytical components of the measurement system, whether the analyses are chemical or biological. This document attempts to provide guidance on both sampling and analytical aspects of biological quality assurance, although sampling design is not addressed in detail. Biological analyses include both structural and functional characterizations of the sample, as defined below.

Structural characterization involves enumeration and identification of life forms (e.g., trees, fish, benthic organisms, plankton, bacteria), while functional characterization involves measurement of a biological response to the sample (e.g., toxicity). Methods of sample collection and characterization differ substantially between aquatic (water, sediment) and terrestrial (air, soil) samples; however, similar quality assurance principles and procedures are applicable.

Biological methods tend to be less standardized than chemical methods. While there is a great need for biological method validation and standardization, as noted by the Alliston Workshop, this document is not intended as a methods manual. Federal and provincial laboratories and working groups, and international standards organizations, with the assistance of universities and environmental consultants, are now working toward standard protocols in many areas of biological testing.

The focus of this document is on the generic elements of biological quality assurance, as derived from current experience with chemical measurement systems. More specific guidelines are also provided on application of QA principles to particular types of biological measurement systems - toxicity testing, microbiological testing, field bioassessment and paleolimnology. Other biological subject areas could be included in future.

Federal, provincial and private laboratories are invited to build on these guidelines in development of their own biological QA plans. Biological QA practice is expected to evolve considerably over the next decade, as specific QA problems with particular biological measurements are recognized and addressed.

2.0 GOOD LABORATORY PRACTICE

A recent OECD draft council decision, which Canada has signed, calls for member countries to establish national Good Laboratory Practice (GLP) compliance monitoring procedures, designate authorities to monitor compliance, and require test facilities to issue a declaration of compliance with each study report (OECD, 1989). The scope of the compliance program (e.g., chemical, biological or both) is to be defined by each member country.

Except for project-specific details which relate directly to project objectives, a good laboratory QA/QC program amounts to Good Laboratory Practice. A 'laboratory' may be broadly defined to include all sampling, analytical and management personnel and facilities, and GLP principles should extend into the field. For many biological measurement systems, the measurements take place in the field, as well as the laboratory, and are, of necessity, highly integrated with sampling activities.

Integration of project planning, sampling and analytical activities is the hallmark of any good measurement system. Thus, GLP principles are applicable throughout the system, and constitute the basic elements of a good QA program. These elements can be categorized as follows (OECD, 1981a; EPA, 1983/87; ASTM, 1988a):

- responsibilities and authorities - for the laboratory and the proposed project;
- data quality objectives - tailored to individual projects and data uses;
- sampling procedures - if sampling is involved in the project;
- analytical procedures - including biological measurements (tests) as well as supporting chemical measurements;
- QA/QC procedures - which may be packaged with the individual test procedures, or may stand alone; and
- documentation procedures - for storage, retrieval, revision and reporting of written methods, test results and supporting data.

Section 3.0 elaborates on each of these basic QA elements, and each may be further subdivided. Subsequent sections focus on application of the general principles in biological laboratory and field activities.

3.0 ELEMENTS OF A QA/QC PROGRAM

3.1 Responsibilities and Authorities

Responsibilities and authorities are an essential first step in any QA/QC program. Whether outlined in a separate Quality Management Plan, or as one of the procedures in a Standard Operating Procedures document, responsibilities for quality assurance functions must be clearly assigned and understood by all staff. If not clearly assigned, these functions are generally assumed to be someone else's responsibility, and are easily sacrificed to budget or schedule constraints. Most data quality problems are ultimately traceable to such poor management practices.

Data Certification

Along with responsibility for quality assurance functions, there must be sufficient authority to make sure that they are performed. For example, a Director cannot override the quality concerns of his QA Officer or Lab Supervisor in order to save money or meet a deadline. A good mechanism of clearly assigning this authority involves a signature system of data certification, whereby certain signatures are required on every data report. Each signature verifies that the responsible individual has performed his or her assigned quality assurance function, and is satisfied with the quality of the data as stated, and with any data analysis and interpretation provided.

Quality Assurance Policy

Allocation of financial resources to quality assurance functions is the Director's responsibility, and this responsibility must be clearly defined in a QA policy statement. Quality assurance cannot be achieved without a financial commitment. A formal policy statement of commitment to quality provides the mandate by which financial resources may be secured for quality assurance, and effectively drives the whole QA system. It also assures technical personnel of support in placing 'quality first'.

Quality Management Plan

Routine management procedures are often separately documented in the form of a Quality Management Plan for the laboratory or organization. This document focuses on

the responsibilities and authorities of personnel from upper management to bench technicians. The Quality Management Plan should always include a clear QA policy statement that emphasizes upper management commitment to and responsibility for the validity and integrity of all data produced. It should also identify a Quality Assurance Officer reporting directly to upper management who is independent of routine laboratory operations, but technically capable of performing the tests, evaluating performance and determining whether data quality specifications are being met. An example of a Quality Management Plan for a laboratory is provided by the Ontario Ministry of the Environment, Laboratory Services Branch (MOE, 1985a). Such plans have also been developed for interagency QA coordination (e.g., LRTAP, 1986; IJC, 1988).

Quality Assurance Plans

The Quality Management Plan will refer to a Quality Assurance Implementation Plan or equivalent Standard Operating Procedures for technical details of implementing each quality assurance function, such as routine quality control procedures, audit procedures and corrective actions. Guidelines for these activities are included in Section 3.5.

A Quality Assurance Project Plan should also exist for individual projects. This will identify project-specific QA activities, personnel, responsibilities and authorities, as well as data quality objectives for the project, and other project details. It will refer to Quality Management Plan and/or other documents for description of routine laboratory QA personnel and procedures. Preparation guidelines are provided by QAMS (1980).

3.2 Data Quality Objectives

The scientific community is increasingly cognizant of the need for clearly defined data quality objectives which are set prior to data collection and which serve as criteria for data acceptability. These objectives are derived from intended data uses (QAMS, 1986). The process of designing programs to collect the 'best data possible' can be successful, but can also be expensive. Moreover, it is not scientifically rigorous and may still fail to produce data suitable for the intended application. If suitable data are not technically achievable, it is better to determine this via the data quality objective process, before money is spent on an impossible task.

Data are ultimately reduced to a quantitative summary statistic (e.g., a mean concentration or mean LC50) for a particular temporal/spatial domain. From the hypothesis tests, decisions or predictions that will be based on this statistic, a total uncertainty tolerance (ΔC) for the value can be defined. This constraint determines the net bias plus imprecision that will be acceptable (Taylor, 1987):

$$\Delta C \geq \text{bias} + CI$$

where CI is the confidence interval of the summary statistic in terms of total imprecision (S) and number of samples included (n). Both bias and precision may vary with response level, and values which pertain near the decision point should be used. Some simple examples are illustrated in Appendix 1.

Data quality objectives can then be set as upper limits for S, for the sampling method and analytical components of S, and/or for the minimum number of samples. If the domain of the summary statistic includes multiple stations or time points, there may also be an environmental component to S; this cannot be controlled by selection of sampling or analytical methods, but should be considered in defining permissible sampling method or analytical errors. Often the environmental variance is so large that there is little to be gained by attempting to constrain sampling method or analytical error, but this must be demonstrated before relaxing the constraints.

Data quality objectives should also be established for limits of analytical quantitation (sensitivity). These limits depend on the intended data use. If data are used only for compliance evaluation, quantitation limits an order of magnitude below the compliance criteria should be sufficient. On the other hand, if data are intended for descriptive monitoring, or predictive modelling, most of the data should be above quantitation limits, and these limits should be well below the expected ambient levels over most of the study area.

3.3 Sampling Procedures

Sampling procedures <u>are</u> of primary importance in most studies, and should not be overlooked, even by laboratories primarily engaged in analysis of samples after collection. Sampling error is usually a large component, and often the dominant

component, of uncertainty in environmental measurements. For example, a 'representative' sample is not easily obtained from a discharge with fluctuating or cyclical output. Continuous, flow-proportional sampling is the best way of dealing with these situations. A single grab sample is usually not adequate for quantitative representation of a dynamic discharge. Optimum grab sampling programs can be designed if the dynamic characteristics of the discharge are known (see, for example, Gilbert, 1987). If they are not known, some preliminary effort to characterize temporal variation must be included in the study design.

Sampling design is ultimately driven by the questions being asked and the specific hypotheses which the Project Manager wishes to test. Simple random or stratified random designs are common (Appendix 1). The Project Manager should ensure that the design is optimal and sufficient to satisfy his stated data quality objectives. This should be a quantitative derivation of the number of samples needed and their spatial/temporal distribution. Some simple examples are given in Appendix 1. Consideration of expected analytical quality (precision, accuracy and sensitivity of analysis), as well as sampling precision and environmental variability, will be part of this evaluation (Section 3.2).

Laboratory Managers should be familiar with the data quality objective process and prepared to assist Project Managers in determining whether analytical quality will be consistent with the overall data quality needs of each study. They may also advise on sampling techniques to reduce and/or permit estimation of sampling error.

The quality assurance principles associated with sampling are, in many respects, similar to analytical QA principles. They involve good housekeeping and good record keeping. Equipment should be regularly maintained, checked for proper operation, calibrated as required (e.g., an automatic sampler), and cleaned as required, to prevent cross-contamination of samples. Sample collection devices are usually rinsed with sample material immediately prior to sample collection. Procedures for doing these things should be included or referenced in Standard Operating Procedure documents, responsibilities for doing these things should be clearly assigned, and log books should be employed to record routine use, maintenance, calibration and cleaning of equipment.

Written procedures for collection and preservation of samples are required, and these should include sampling forms, log books, labelling and numbering systems, and shipping

procedures. If samples are for possible litigation, chain-of-custody documentation must be used. Procedures may call for quality control samples, such as trip blanks, equipment rinsates (i.e., rinse material for analysis), and field duplicates to be analyzed, in accordance with data quality needs (Section 3.2).

3.4 Analytical Procedures

Analytical and biological test procedures are subject to the same degree of documentation as sampling procedures. Chain-of-custody documentation continues into the laboratory, where records should permit sample-specific retrieval of sample receipt information, analytical results if completed, and associated quality control data at any time. Procedures may call for control samples such as method blanks, spiked blanks, matrix spikes and lab replicates (split samples) to be analyzed, in accordance with data quality needs (Section 3.2).

Laboratory Facilities

Most biological and supporting chemical measurements are performed in a laboratory. The layout of the laboratory facility is important in preventing mix-up or cross-contamination of samples. There should be separate defined areas for receiving, reagent storage, sample storage, sample preparation and sample testing. Samples should be logged into the data storage system on arrival by client name, sample name, number and date. They should be labelled and scheduled for analysis immediately, the schedule depending on permissible holding times as per Standard Operating Procedures. Cold storage is often required to prolong the life of the sample and maintain sample integrity.

New reagents and instruments should be logged and labelled in similar fashion with expiry or maintenance dates recorded on labels. Instrument manuals, calibration procedures and log books should be filed near each instrument, and also in a central location. Working solutions should be stored near their use areas, under appropriate conditions, and labelled with preparation and expiry dates.

The sample preparation area should be separate from the sample testing area. In the preparation area, such activities as extraction or toxicity reduction treatments would be performed. If hazardous materials are tested, a fume hood and safety equipment should

be available in this area. Near the sample preparation area, separate areas for clean and dirty glassware should be defined. Both sample preparation and testing areas should have adequate ventilation to permit rapid clearing of volatile releases. Copies of test methods should be available near the testing area. Bound (or electronic) bench sheets for recording test data should be available in the sample testing area.

3.5 Quality Assurance Procedures

Quality assurance procedures should be included in Standard Operating Procedure documents, either as part of each sampling and analytical procedure, or as a separate procedural system, or both. QC procedures will call for analysis of quality control samples of various types with a specified frequency, and will specify acceptance criteria for each batch of samples analyzed, based on the results of control sample analyses. Both sampling and analytical control samples should be included. The QA Officer will review these QC results on a regular basis as part of his periodic QA assessments, as per defined QA procedures.

Audits and QA Reports

Quality assurance assessment procedures should include the following audit activities:

- o review of QC sample results;
- o performance assessment by submission of 'blind' QC samples (both composition, and QC status if possible, unknown to analyst);
- o knowledge testing and training sessions;
- o review of data prior to release, with checks for logical inconsistencies, calculation errors and transcription errors;
- o review of other lab records for adequate documentation of samples, reagents, instruments, methods and method changes; and
- o participation in interlaboratory comparison and accreditation studies.

Reports should be prepared on a defined schedule, summarizing each quality assurance assessment. These reports should be prepared and signed by the QA Officer, reviewed and signed by the Laboratory Manager, and filed as part of the long-term quality assurance record of the laboratory. Reports should clearly identify any quality problems identified and corrective actions taken.

Sampling Quality Control

The total precision of measurement is quantified by inclusion of field replicates in the sampling design. The entire sample collection is replicated at essentially the same point in time and space. The variance (S_m^2) of measurements on these replicates provides an inverse measure of precision (e.g., standard deviation S, or relative standard deviation = coefficient of variation $S/\bar{x}$).

The number of samples to be replicated in the field, and the number of replicates per sample, are elements of sampling design to be derived from the intended data uses and overall uncertainty tolerance of each use (Section 3.2, Appendix 1).

Random selection of sampling units to be replicated is desireable (if all units are not replicated). The proportion of samples replicated may vary among sub-strata of the population under study, according to the sampling design. Field replicates should not be construed as primary sampling units; they are used in estimation of measurement error, but are not pooled indiscriminately with samples from other stations for description of a larger area.

Precision of sample collection can be inferred by subtracting analytical variance (S_a^2 from the total measurement variance (S_m^2) of field replicate samples to obtain the sampling method variance (S_s^2). Analytical variance is determined from measurements on lab replicates, prepared in the laboratory by homogeneous splitting of a single field sample (page 3.8). There may also be spatial or temporal environmental variance (S_e^2) between the stations or time points of a study. Total variance (S^2) includes S_m^2 plus S_e^2.

Sampling bias is difficult to quantify without a complete prior characterization of the study area, in which case a study is unnecessary. However, sampling bias can be avoided by following a formal sampling design (e.g., random, stratified random, systematic) on which subsequent data manipulations will be based. The design must define the sampling units and the spatial/temporal boundaries of the study.

Contamination during the sampling process can be controlled with trip blanks and equipment rinsates. Trip blanks are prepared in the laboratory, taken to the field, prepared as regular samples and returned to the laboratory for analysis. Rinsates are

water or solvent samples processed through the sample collection device(s) in the field just prior to sampling. Such samples provide a 'background' chemical concentration or biological response.

The use of control sample results in determining the validity of the sampling effort, and acceptance of concurrent routine sample data, should be clearly specified in project plans prior to sample collection.

Analytical Quality Control

Analytical QC samples should be tested frequently as outlined in Standard Operating Procedure documents. These usually include method blanks, lab replicates and standards of some type with each batch of samples tested. Blanks are prepared from 'contaminant-free' matrix, but include any reagents used in analytical processing of actual samples. They provide a check against inadvertent contamination of samples in the laboratory during sample preparation and analysis. They are not expected to give a significant test response. In chemical analysis, the variability of response to blank samples, or blanks spiked at very low levels, is used to define the Limit of Detection of the test (EPA, 1984; ASTM, 1985; MOE, 1988a). A blank response above this limit is indicative of contamination. When this happens, the actual samples in the same test batch should be re-analyzed, if possible, otherwise the blank response must be reported with the associated test results. Most authors recommend against blank-correction of test data (Hunt and Wilson, 1986); however, current practice varies, and any correction performed must be stated.

Lab replicates are samples split in the laboratory (with careful attention to homogenation of sample prior to splitting) and separately analyzed to obtain an estimate of test variability (i.e., analytical precision). Analytical precision can be expressed in different ways, but is usually expressed as the standard deviation (S_a) of lab duplicate responses (ASTM, 1985). A small standard deviation represents good precision. Precision is almost always a function of response strength - as the response increases so does the standard deviation. Consequently, some authors prefer to speak of the Relative Standard Deviation or Coefficient of Variation, which is simply S expressed as a percentage of mean response. This quantity is more stable, but still tends to vary with response, particularly at low response levels. Thus, while a precision function may be defined for a

given sample matrix, it is difficult to quote a single value as the precision of a test method. An approximate value may be quoted for a specific response range, but the range, units and matrix should always be stated.

Standards of various types are used to assess accuracy of chemical measurement. These may be Standard Reference Materials (certified as such by a recognized agency - e.g., National Bureau of Standards, Environment Canada, National Research Council of Canada - after repeated analysis) or Spikes. Spiked samples are prepared by each analytical laboratory by adding known amounts of analyte (or a surrogate) to a natural sample (matrix spike) or a blank. Standard Reference Materials are usually prepared in a simpler matrix than that of actual samples, and are therefore often easier to analyze accurately. Matrix spikes are more representative of performance on actual samples, but the true concentrations of spikes are more open to question than those of Standard Reference Materials. In either case, the measurement bias is usually quantified as the average deviation (positive or negative) from true value, expressed as a percentage of the true value. Recovery is simply 100% minus this percent deviation, and may be greater or less than 100%. Accuracy is 100% minus the absolute value of percent deviation on a single measurement, and never exceeds 100%. This term is sometimes reserved for the results of Standard Reference Materials analysis; however, Standard Reference Materials do not exist for biological response measurements (see Section 4.4).

Control charts (or tables) should always be used to monitor analytical precision and accuracy in the laboratory. Many types of control charts are available. Techniques for chemical analysis reviewed by ASTM (1986) and Dux (1986) are also generally applicable to biological analysis. They fall into two categories - average and range charts. Average charts monitor accuracy (or recovery), while range charts monitor precision. Examples are illustrated in Figures 3.1 and 3.2. When the short-term average value of a standard sample falls outside its normal range, an accuracy problem is indicated. When the range between replicates becomes abnormally large, a precision problem is indicated. There may be a series of charts for each measurement system, according to sample (matrix) type and concentration (response) categories. Acceptance limits on these charts are updated on a regular basis, but tend to stabilize once a large body of control data has been accumulated.

Action guidelines for control chart exceedences should be in place. There should be a chart entry for each batch of samples concurrently tested. Actions may include re-testing the samples in an out-of-control batch, or reporting results with appropriate data quality warnings.

Corrective Actions

Quality assurance plans must include clearly defined procedures to trigger for specific corrective action. A feedback system is essential to a functioning QA program. Problems identified in audits or by end users of the reported data must be investigated and rectified. Common triggers and initial corrective actions include rejection of data generated during times at which the test method is known or suspected to be out of statistical control, rejection of animal test data if control responses exceed previously-established tolerances, laboratory managers' refusal to accept samples which have not been collected, identified, preserved, transported or handled in accord with previously-established Standard Operating Procedures, etc. Specific follow-up corrective measures will then be required to ensure problems are rectified and subsequent data are of suitable quality.

Authority for the most drastic corrective actions is normally vested in Project Managers and QA officers, and good QA plans will define the limits of authority for all specified corrective actions. Failure to establish data acceptability criteria and to provide for remedial action appears to be one of the most common oversights in newly-established QA programs.

3.6 Documentation Procedures

Documentation procedures should also be included in Standard Operating Procedure manuals, to ensure adequate and consistent documentation of field and laboratory methods, quality assurance and analytical results. The contents and location of each file should be specified, and access should be controlled by an individual who is assigned this responsibility. Often this individual is the QA officer. Retention times should be clearly defined, along with procedures for revision and disposal of documents. All documents require dates and revision numbers so that new and old versions cannot be confused.

Analytical or biological test results are often stored electronically in computer files. This facilitates data review and checking; however, the same principles of access, revision and retention should apply. Moreover, there should always be some degree of human verification of computer software to guard against inadvertent program changes, and there should always be back-up systems to permit recovery of data lost due to hard-disk failures or operator errors. File security systems (e.g., 'read only' files) can be used to control access and maintain file integrity.

Data Reporting and Handling

Guidelines for reporting of data should also be in place. Data should be reported to the correct number of significant figures. They should always be reported with some indication of analytical sensitivity, precision and, if possible, accuracy (each measure defined), and there should be a description of, or reference to, the analytical methods used. Data are generally not either blank-corrected or recovery-corrected prior to reporting; however, some instruments and computer programs automatically do this, and any such correction should be stated.

Reporting of low-level data varies among laboratories. Many laboratories do not report a quantitative value below some reporting limit. Results are simply reported as 'less than' the reporting limit, or 'non-toxic' in toxicity assessments. In these cases, the reporting limit must be stated. ASTM (1985) recommends that all values be reported, even if there is a non-significant or negative response. Chemical concentrations lower than the Criterion of Detection (Figure 3.3) should be reported with a 'T' flag for 'trace'. If there is no test response because the test has a response threshold, the threshold is reported with a 'W' flag. Project Managers should be aware of, and may wish to specify, reporting methods for a project.

Trace data (T flagged) should always be used in calculation of summary statistics and trends. Special methods of mean and variance estimation are recommended for inclusion of censored samples (W flagged) in statistical summaries (Helsel, 1986; Gilbert, 1987). Exclusion of censored data is not recommended, either for mean or trend estimation.

FIGURE 3.1
Range and Average Charts for Control of
Precision and Accuracy (After EPA, 1979)

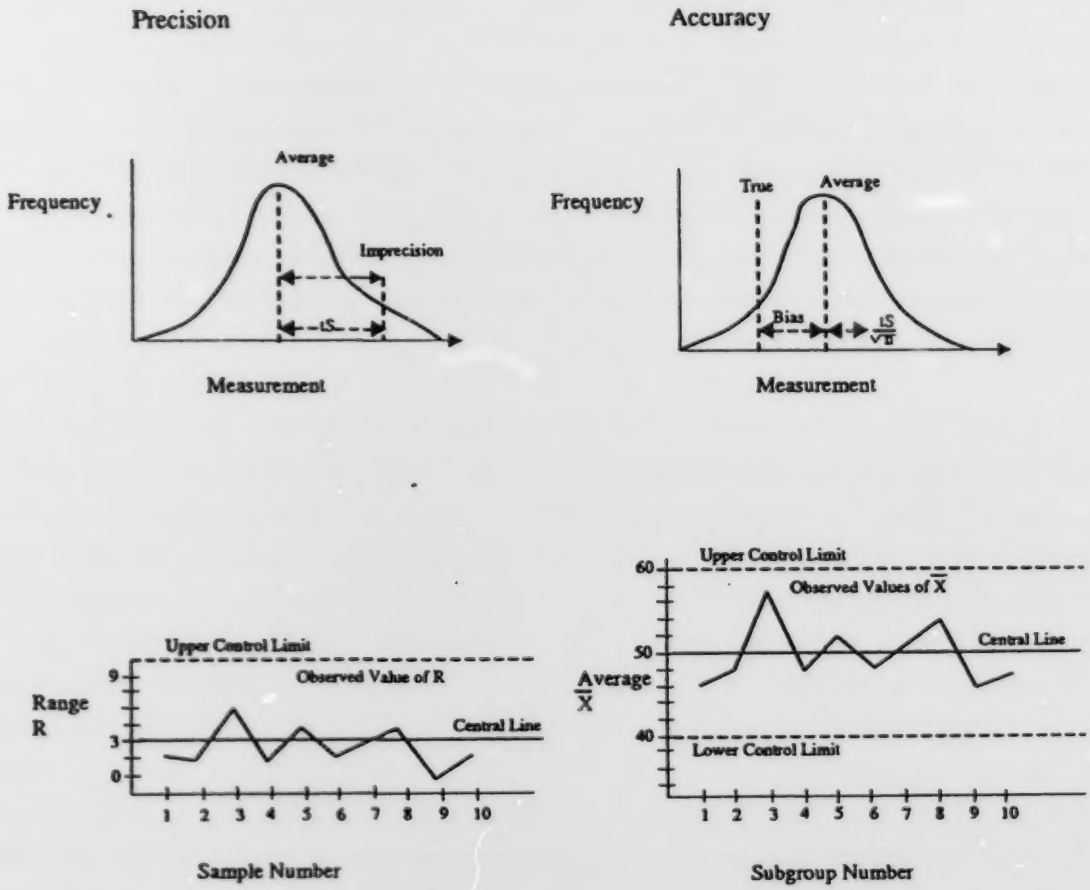


FIGURE 3.2
Range Chart for Response-Dependent Precision Control (After MOE, 1984)

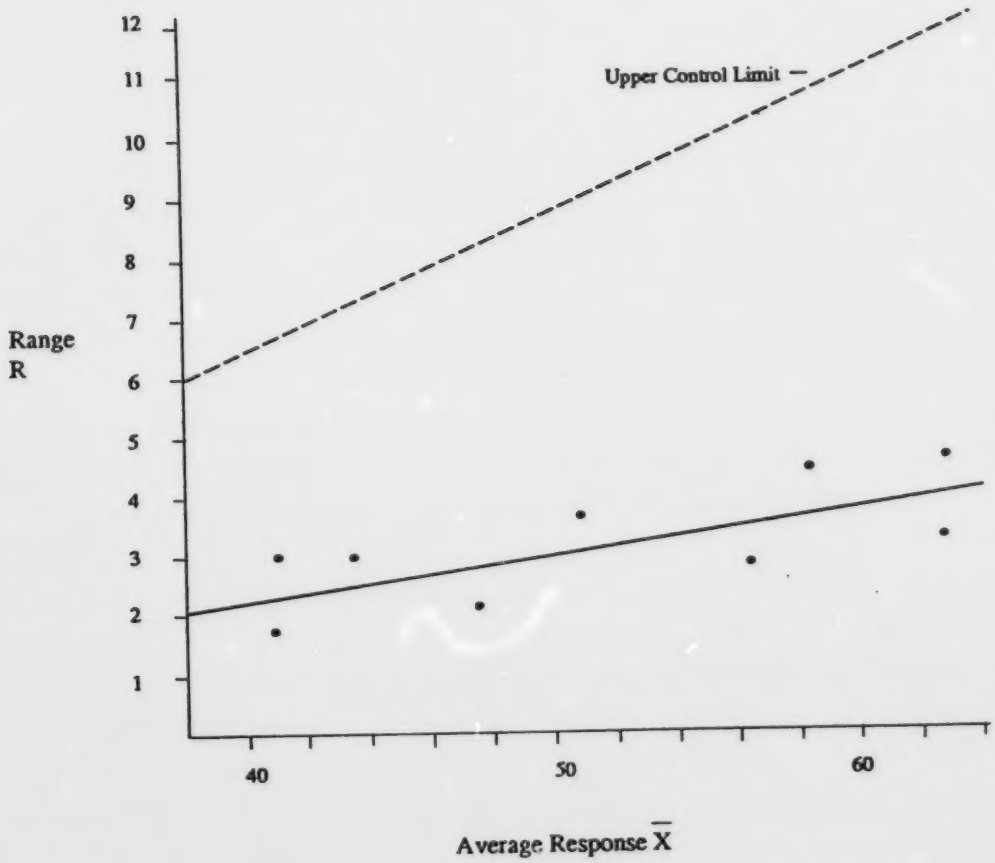
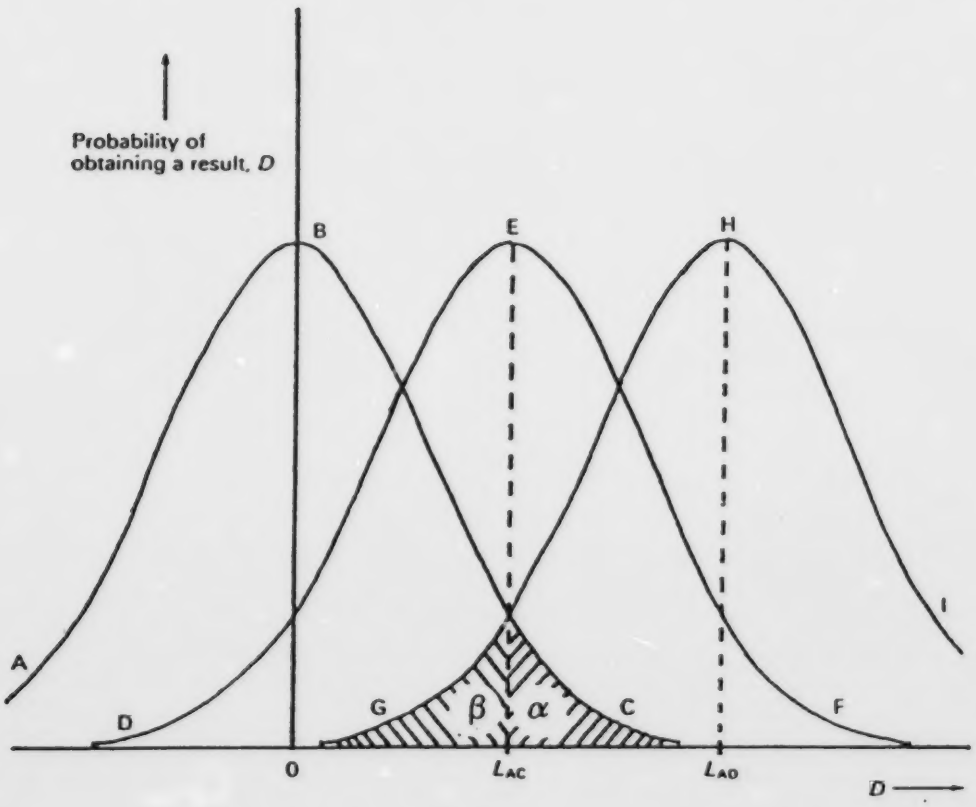


FIGURE 3.3: CRITERION AND LIMIT OF DETECTION (after Hunt and Wilson, 1986)



L_{AC} = Criterion of Detection

L_{AD} = Limit of Detection

$\alpha = \beta = 0.05$

4.0 QA IN AQUATIC TOXICITY TESTING

Toxicity measures may be considered as additional properties of the sample to be analyzed, using the test organism as the analytical instrument to measure a property of the sample (i.e., a toxicity endpoint). Consequently, the general elements of quality assurance outlined in Section 3.0 also apply in the toxicity laboratory (MacGregor and Doe, 1987). In addition to these general QA elements, there are unique sources of error and corresponding quality control checks to be performed. This section focuses on those unique aspects of toxicity testing.

A variety of toxicity testing approaches are used, ranging from laboratory to in situ settings, and from single-species to multi-species assemblages of test organisms. Exposures may be acute (short duration) or chronic (long-term); endpoints may include lethality, sublethal organism responses, or various community structural and functional responses. Results may be expressed in terms of sample concentration required to induce the response in 50% of test organisms (or communities), or as the percent of organisms responding at a single concentration. While the examples of this section are drawn primarily from single-species laboratory testing experience, the QA principles and approaches outlined may be considered generally applicable.

4.1 Sample Collection and Storage

The volume of sample needed for each test should be defined in Standard Operating Procedures and QA plans, along with any constraints on the collection vessel and shipping procedures appropriate for different types of samples. For example, headspace should be eliminated from the sample container at the time of collection, to limit aerobic degradation of the sample. For samples to be used in litigation, the sample collection, shipping and receipt at the laboratory should be documented with dates and signatures on chain-of-custody forms. Sample acceptance criteria may be defined in terms of age and temperature on arrival, headspace, condition of custody seals, and/or sample volume.

Samples for toxicity testing are not chemically preserved, and their stability characteristics are often unknown. Therefore, holding time prior to testing should be minimized, and specified in test protocols (e.g., five-day maximum for acute trout and Daphnia tests (EPS, 1980; MOE, 1983, 1988b); three-day maximum for acute trout test

(Environment Canada, 1989a)). Refrigeration (not freezing) is recommended throughout the holding period, but samples must be brought to the standard test temperature prior to testing. Sediments represent a particular problem with respect to maintaining *in situ* properties for sediment toxicity testing (see ASTM, 1989a).

4.2 Records of Test Organisms

Records must be kept on the history of test organisms, just as records are kept on each analytical instrument. Animal records should include the quantity and source of each shipment (supplier or collection location), date of arrival, verification of taxonomic identity, nature and date of health checks performed, any health certificates, and details of the acclimation history. All animals must be acclimated to new holding conditions before use, so that the stress of acclimation does not influence test results. Acclimation periods should be specified in test protocols (e.g., two-week minimum for acute trout, *Daphnia* and stickleback tests, Environment Canada, 1989a,b,c). Generally, holding conditions are the same as test conditions with respect to critical factors such as temperature, light and photoperiod. Ideally, water supply is also the same, except for toxicant effects during testing (see below). Some aquatic tests allow for a special synthetic dilution water, but there should be acclimation to such water before testing. Physiological measurements indicative of general stress may be used to confirm that acclimation to a new condition has occurred (e.g., plasma glucose, plasma protein; see Bouck, 1984); however, baseline data are required, and sampling must be rapid to avoid measuring capture/sampling stress.

Organism records should include any death or disease occurrence, disease treatment or preventative treatment. Diseased organisms should always be quarantined from the main holding facility. In most toxicity testing laboratories, diseased animals are sacrificed immediately, and any organisms recovered from disease or test exposure are not used for further testing. All handling equipment should be dedicated for use on either healthy organisms prior to testing, organisms during testing, or unhealthy organisms. Fish holding tanks and associated handling equipment should be disinfected between batches of fish.

Regular schedules for feeding and cleaning of organism enclosures should be established. These procedures, along with record keeping, acclimation, health care and

test procedures, should be included in Standard Operating Procedure documents. Acceptance criteria for food and water supplies, and other environmental conditions such as temperature, pH, light, dissolved O_2 in water, water change rates and holding densities, should also be established and compliance should be monitored on a regular basis. Food supplies may be subject to minimum nutritional and maximum contaminant specifications. Criteria specified in test protocols should be considered in establishing laboratory criteria. Examples of criteria for reconstituted waters and other dilution waters for aquatic toxicity testing are listed in Appendix 1 (Table A4.1).

Water supply treatment by activated carbon filtration, particle filtration, ultraviolet irradiation, ion exchange or reverse osmosis filtration may be required to meet dilution water standards. Water treatments and their rationale should always be identified in Standard Operating Procedures, and maintenance of acceptable water quality demonstrated by regular chemical monitoring.

4.3 Test Methods

Test methods should be described in Standard Operating Procedure documents and referenced to internationally recognized test methods (e.g., OECD, 1981b; ISO, 1982, 1984a; APHA, 1985; FAO, 1987; ASTM, 1988b,c, 1989b,c). Published test methods of federal and provincial/state jurisdictions are also acceptable (e.g., Alberta, 1978; EPS, 1980; B.C. MOE, 1982; MOE, 1983, 1988b; EPA, 1985a,b; NYDEC, 1985; Environment Canada, 1989a,b,c). Methods should always specify acclimation periods, dose level selection, dosing schedules, routes of administration, test duration, test endpoint, species, lifestage, age, and/or size of test organisms and test acceptability criteria. In aquatic toxicity testing, water change rates, loading rates (organism mass/litre or mass/litre/day) and aeration rates should also be described. Most published methods specify these criteria, along with acceptable environmental conditions such as light regime, temperature, dilution water quality and test water quality. In sediment tests, establishment of water/sediment equilibrium is essential. All these critical test conditions should be measured and recorded with the test result. Sometimes test solutions are modified (by gentle aeration or pH adjustment) to meet test water quality criteria for dissolved oxygen and pH; such treatments are also recorded and reported with test results.

Aquatic toxicity tests may be performed in static, semi-static (batch water changes) or continuous-flow mode, and the mode of operation should always be stated. Toxicant delivery systems (diluters) for continuous-flow tests are elaborate and require calibration. Test methods should include a description of calibration procedures and precision and frequency of calibration in continuous-flow tests. It has been recommended that diluter flow rates should not vary by more than 10% among test chambers (EPA, 1979).

Exposure verification by chemical analysis during the test is always desirable, but particularly so in continuous-flow tests. When the toxicant is a mixture, conductivity and/or dissolved organic carbon often provide good surrogates for chemical analysis. It has been recommended that test concentrations in each chamber should not vary more than 5% over the test duration (EPA, 1979). Verification frequencies, criteria for acceptance of discrepancies, and action guidelines for excessive discrepancies should be included in test methods.

For toxicants of limited solubility, a carrier solvent is often utilized. However, the solvents themselves can be toxic, or may influence partitioning and uptake rates of compounds into tissues. It has been recommended that carrier concentrations should not exceed absolute concentrations of 0.1 g/L in test solutions (EPA, 1978). Moreover, carriers should never exceed 10% of their own LC50. Preliminary experiments with different carrier concentrations at a fixed toxicant concentration are advisable, to determine no-effect carrier concentrations, and ensure that test concentrations of carrier are below these levels. When carriers are used, there should be a solvent control in addition to the usual dilution water control, and these two controls should not differ significantly. Generally, use of carrier solvents should be minimized.

Range tests of short duration (less than 24 h), but wide concentration range (single or half orders of magnitude), are often used to ensure optimum selection of dose levels for the definitive toxicity test. Dose level spacing in the definitive test should comply with test protocols (arithmetic or geometric series) and concentrations should bracket the true median lethal (LC50) concentration. Ideally, there should be at least one partial response both above and below, or equal to, the 50% response level. If this is not the case, any LC50 reported should be flagged as an extrapolation. The EPA (1979) and Environment Canada (1989a,b,c) recommend at least five dose levels, plus a control, with

partial responses above 65% and below 35%. However, for some toxicants with a steep concentration-response curve, this dose level spacing may approach the range of error in solution preparation. An LC50 can be calculated with no partial responses, but will have wider confidence limits. Ultimately, the data quality needs of the study (i.e., maximum uncertainty permitted) will place limits on dose level spacing.

Feeding is usually not recommended during acute aquatic toxicity tests (up to 96 h). However, it is necessary during chronic (long-term) tests. Recommendations on feeding are provided by Lanno (1989). The feeding regime should be described in test methods, and should be uniform among dose levels. Provision must also be made for prompt removal of unused food and dead test organisms that could otherwise foul the test solutions.

4.0 Quality Control

Control Response

Most acute test protocols require a control response of less than or equal to 10% for the test to be valid. This is analogous to the blank acceptance criterion in chemical analysis. If greater test sensitivity is required, a smaller percentage may be set as an acceptance criterion; however, the ability to detect a specified control response depends on the number of organisms in the control group. For example, ten organisms are needed to detect a 10% response, while 20 are needed to detect a 5% response. OECD permits as few as seven organisms, which implies a detection limit of 15% response. Following the ASTM recommendation for chemical analysis, test responses at each concentration should be reported even if they do not exceed control responses or permit an LC50 (EC50) calculation.

For some test endpoints (e.g., algal growth inhibition), the control is used as a reference value, and the control always exhibits 0% inhibition relative to itself. This is analogous to blank correction in chemical analysis, although the correction formula is different. Blank correction can also be applied to mortality or other quantal response data using Abbott's correction formula (Finney, 1981). Correction has little effect and is probably unnecessary for a response that should be rare in a valid control, but may be needed in other cases. Any data correction should be reported with the test result.

In sediment toxicity testing, controls may either be the same clean sediments that are spiked to produce the test exposure levels, or reference sediments comparable to the test sample(s) in all respects except toxicant contamination (e.g., ASTM, 1989b,c). Reference controls must be carefully matched to test samples with respect to particle size and TOC, both of which may strongly affect contaminant partitioning between water and sediment.

Replicate Tests

Simultaneous replicate tests should be performed routinely, as in chemical analysis, in order to quantify the within-lab precision or reproducibility of the test procedure on selected samples of different types. These samples should be selected from the regular sample stream, and may also include reference toxicants. Criteria for acceptance of tests based on within-lab agreement of replicate LC50 (or other endpoint) determinations should be established. Range-type control charts can be used, as in the chemistry laboratory (Figure 4.1). Another simple control device is a plot of high vs. low duplicate LC50 values. If the slope of the line fit to this plot is 1, then the intercept (a) represents within-laboratory test variation ($S = a / \sqrt{2}$). If the slope is not 1, then there is a systematic change in reproducibility with changing toxicity. In either case, an upper control limit can be established as a multiple of R, the vertical distance from regression line to diagonal (Figure 4.1).

In control charting, it is assumed that deviations from the central line are normally distributed, and that the central line is linear. These assumptions should be verified by examination of the data, and normalizing transformations can be used as appropriate. If desired, separate charts can be produced for different sample types and response ranges. Outliers are usually eliminated before computation of control limits. This is an iterative procedure (ASTM (D4210), 1985).

Preparation of replicate dose series should be entirely separate, since this preparation is one component of within-laboratory variation. Replicate series may, in fact, use offset dose levels to encompass variation associated with dose selection. Each LC50 will have its own confidence limits, as defined by the calculation procedure. Various calculation methods are available (e.g., probit, binomial, moving average; see Stephan, 1977). These confidence limits reflect number of organisms per dose level, variation in response

between dose levels, and the number and spacing of dose levels that exhibit partial responses. They are affected by many of the same factors that influence agreement between replicate tests, including dilution technique, organism handling, and organism allocation to test vessels. Thus, replicate *LC50* values should never be significantly different, and should always have overlapping confidence intervals.

Results of replicate tests should be individually reported. If a single measure of toxicity is also desired, the unweighted mean LC50 is not necessarily the best estimate. Hong et al. (1988) describe a maximum likelihood procedure for combining replicate probit lines (if not significantly different) to derive a single mean and confidence interval.

Reference Toxicants

Reference toxicants in toxicity testing are the nearest equivalent to the use of chemical standards in chemical analysis. As they are not certified, 'true' responses cannot be cited. Responses are expected to vary with organism stock and dilution water supply. Nevertheless, they provide a reference point against which any changes in the test system can be detected. The LC50 of pentachlorophenate, sodium chloride, or sodium dodecyl sulphate, for example, can be determined on a regular basis, and charted on an average-type control chart, just as an Standard Reference Material concentration would be charted in the chemistry laboratory (Figure 4.2). At a minimum, reference toxicant tests should be performed whenever a new stock of test organisms is acquired, since differences in sensitivity between stocks can bias test results. Recent draft test protocols developed by Environment Canada (1989a, b, c) call for reference toxicant tests at least monthly. Over the long-term, control limits are defined (e.g., $\pm$ 2 S) and any reference test results outside these limits signify an unusual and possibly unacceptable bias in the toxicity testing system. A recent Environment Canada (BEAK, 1990) manual on reference toxicant testing elaborates on control technique, and also recommends an upper limit on the acceptable standard deviation S (coefficient of variation $S/\bar{x}$ less than 30%).

4.5 Interlaboratory Comparison

Interlaboratory comparison is an essential part of any quality assurance program. It is particularly important in toxicity testing since certified standards do not exist, and since

interlaboratory differences in test sensitivity related to dilution waters and animal stocks are expected. Either reference toxicant solutions or toxic effluent samples can be split among laboratories, with careful attention to homogeneity of subsamples, and the results can be used to estimate interlaboratory test variation (e.g., Grothe and Kimerle, 1985; PACE, 1989). With a sufficient number of participating laboratories, outlier tests can also be performed to identify laboratories which consistently report outside the normal range of interlaboratory variation. Outlier data are not necessarily wrong, particularly in cases where methodological flexibility is permitted within existing test protocols. Outlier data should be more closely scrutinized when standard protocols for regulatory use are in effect. Government laboratories or private laboratory associations should take the lead in initiation of interlaboratory studies and follow-up investigations. Participant fees can be used to finance studies.

Demonstration of successful performance in interlaboratory studies is one of many criteria which can be used in laboratory evaluation (Section 4.10). However, success can only be defined in terms of statistical agreement with other laboratories in a study with a sufficient number of participants. Youden (1975) provides guidance on statistical evaluation of interlaboratory studies.

Method Validation

Interlaboratory calibration studies can be used as part of the validation procedure for new test methods or method modifications. Performance criteria should be defined for each method, and may include such factors as within-laboratory and between-laboratory precision. Before any method modification is implemented, a comparison with the original method in terms of the specified performance criteria should be documented.

4.6 Toxicity Degradation Studies

In toxicity degradation studies, a liquid sample is treated (or not) and assessed repeatedly at intervals over time in order to determine the rate and extent of toxicity reduction. Toxicity may be reduced by natural chemical decomposition of toxic agents, volatilization of toxic agents, bacterial degradation of toxic agents, and/or some other artificial treatment process. In either case, the temporal consistency of the test system is particularly critical in interpretation of study results.

The temporal spacing of tests is important when the rate of toxicity reduction is to be studied, and requires some prior knowledge of approximate toxicity reduction rates. If these rates are unknown initially, the first few tests should be very closely spaced in time (e.g., daily) in order to ensure that toxicity reduction is not completed between the first and second tests. The testing interval may be increased as the study progresses, and the time course of toxicity reduction becomes more apparent. There should be at least five tests over the course of the study to permit estimation of the shape of the toxicity reduction curve, and the study should continue until toxicity has been reduced to 10% of its original value. In chemical biodegradation studies, the ISO defines 'degradation time' as the time from 90% to 10% degradation (ISO, 1984b). To document temporal consistency in the test system, control samples should be tested throughout the study. Control samples include the dilution water controls (i.e., blanks) that are part of every test, replicate tests and reference toxicant tests. Replicate tests need not be performed at every time point, but should, at a minimum, be performed near the start, middle and end of the study. The 'time-zero' parent compound may be used as the reference toxicant in degradation studies, if its stability can be maintained (e.g., by cold storage). Reference toxicants should be tested at every time point.

In biodegradation studies, bacteria and nutrient media may be added to the sample, prior to toxicity testing of the sample at subsequent intervals. Several additional controls should be included as treatments in the experimental design. A typical design would be as follows:

- media control - media seeded with bacteria;
- sterile test - sterilized media with toxicant; and
- bacterial test - seeded media with toxicant.

The media control verifies that neither the media nor media metabolites are toxic to test organisms, and that the bacterial seed does not contain pathogens. Toxicity tests on this sample should never give a toxic response. The sterile test series will indicate how much toxicity reduction over time is attributable strictly to chemical degradation of toxicant. The bacterial test will reflect both bacterial and chemical degradation.

The media control is particularly important. For some toxicants, a period of increasing toxicity is observed when bacteria form metabolites that are more toxic than the original

test compound. In such cases, it is necessary to demonstrate that media metabolites are not responsible.

Chemical analysis is desirable in toxicity reduction studies in order to provide a chemical explanation for the observed toxicity trend. In an organic mixture, dissolved organic carbon is a good surrogate for chemical analysis. Reduction in this constituent indicates degradation of the toxicant, even when toxicity increases as a metabolite effect.

Toxicity Identification/Reduction Evaluation

A system of artificial treatments aimed at identification of the toxicant(s) primarily responsible for sample toxicity has been outlined by the U.S. EPA (Fava et al., 1988; Mount, 1988; Botts et al., 1989). In these procedures, toxicity is tested before and after various chemical fractionation and chemical detoxification treatments. The temporal rate of toxicity reduction is not of primary interest. Some QA/QC recommendations are included in the protocol documents; however, the principles outlined in preceding sections of this guideline document are generally applicable.

4.7 Raw Sample Compliance Tests

In regulatory applications of toxicity data, compliance criteria are often applied to single undiluted samples. For example, it may be required that an effluent be "non-toxic at all times"; any sample exhibiting toxicity represents a violation. Various definitions have been used in practice to recognize 'non-toxic' samples, and these have not always been clearly stated. Survival of 80% or more of test organisms in an undiluted sample is used by the U.S. EPA in its National Pollution Discharge Elimination System as a criterion of non-toxicity. Environment Canada (EPS, 1980) has indicated that a sample may be considered acutely non-lethal if more than 50% of test fish survive.

A percent survival estimate has an associated confidence interval, just as an LC50 or EC50 estimate has an associated confidence interval. The compliance criterion must lie outside the confidence interval before a violation can be inferred with reasonable confidence, based on a single sample. Binomial confidence limits for single sample survival estimates depend only on survival and number of test organisms used, and will be very wide when few test organisms are exposed (Dorn and Rodgers, 1989). For example, with ten test organisms, 50% survival could not be distinguished from an 80% criterion.

A raw sample survival estimation system should be subject to the same quality control checks as any other toxicity measurement system. These may include range and average control charting procedures, and use of reference toxicants. Dilution water controls with test acceptance criteria are also appropriate as a check on health of test organisms, even though no sample dilution takes place. As with LC50 endpoints, two concurrent replicate tests should not differ significantly. Raw sample test protocols should specify a sufficient number of test organisms to achieve acceptable precision in the response range of the compliance criterion.

4.8 Multi-species Toxicity Tests

Bioassay systems involving simultaneous exposure of multiple species are more realistic than single-species tests for representation of probable ecosystem responses to toxic discharge. In natural aquatic communities, species interactions are important. For example, toxic effects need not reduce all susceptible populations; an affected population can increase if competing species are even more susceptible to the toxicant. Microcosm and mesocosm tests allow for such interactions.

Microcosms are smaller, more artificial and less realistic than mesocosms, but tend to be more replicable (Giesy and Allred, 1985). While more complex than single-species tests, they are simpler than mesocosms. Microcosms attempt to represent key structural or functional ecosystem components, to permit ecologically valid conclusions for a general type of community, whereas mesocosms attempt to simulate a specific community. Moreover, microcosm protocols attempt to standardize environmental conditions and fluctuations, whereas mesocosm tests attempt to represent site-specific environmental variability. Microcosms are generally set up in the laboratory, while mesocosms are typically large on-site enclosures (e.g., the MERL mesocosm, Santschi, 1985).

Standard protocols for aquatic microcosms have been developed (ASTM, 1988d), but are flexible as to particular endpoints measured. Typical endpoints include toxicant (or metabolite) bioaccumulation, species diversity, biomass, trophic structure (e.g., % consumers or producers), production/respiration ratios, population age structure (e.g., % juvenile, adult, post-reproductive), population density, reproduction, mortality or growth. Toxicant degradation and persistence are often measured simultaneously. The ASTM test is sixty-three days in duration.

Quality control principles applicable to single-species testing are also applicable here. In addition, some unique controls are appropriate. For example, nutrient controls are recommended with any test in which the toxicant has a potential nutritional effect on some organisms, since nutrient cycling often drives corresponding population cycles and shifts in species dominance. A nutrient control has similar nutrient value as the toxicant exposure, but is not toxic to any of the test organisms.

Reinoculation of organisms into the test system is performed repeatedly over the course of the test, in order to simulate the immigration into an exposed area that would continually occur in a real ecosystem. It is important that reinoculation densities are below the detection level of the organism enumeration protocol, and that reinoculations occur immediately after (not before) sampling for enumeration.

Replication is recommended with every microcosm test, since the approach involves statistical comparison of treatment and control measurements through time, and since each test may have different contaminating microorganisms present. Control microcosms, like natural ecosystems, are intrinsically quite variable, and it is important to address whether apparent treatment responses are within this range of natural variability at any point in time. The degree of replication required will depend upon the endpoint of interest, its natural variability, and the resolution desired of the test (e.g., the smallest treatment-control difference, as a fraction of standard deviation, considered essential to detect).

Control acceptance criteria should be defined prior to testing, as in single-species toxicity tests. However, the criteria will be more elaborate in microcosm or mesocosm studies. For example, ASTM (1988d) defines an acceptable range for peak algal volume and peak daphnid density, as well as nutrient depletion in the first month, and daytime oxygen gain. The range of normal performance will be specific to each test protocol.

The specific information to be documented as part of each test should also be defined in protocols. In general, this information must be sufficient to demonstrate test validity according to all acceptance criteria, should include individual replicate responses as well as summary statistics for each treatment, and will describe any deviation from protocol or statistical assumptions violated. Statistical summary of the control should include the calculated Interval of Non-significant Difference (based on control variability) at each measurement time point.

4.9 *In Situ* Bioassays

Toxicity tests performed in the field with caged organisms are similar in principle to laboratory bioassays. However, test conditions are less subject to experimental control, and test results tend therefore to be more variable with greater associated uncertainty. A typical situation might involve cages placed in receiving water at different distances from a known or suspected toxicant source.

In such situations, both the toxicant (e.g., effluent) exposure, and the dilution (i.e., receiving) water quality may be variable. In the laboratory, dose verification and documentation of dilution water quality are recommended. In a field situation, the corresponding chemical measurements must be frequently performed over the course of the test in order to document the temporal variation as well as the average exposure conditions at each treatment level (i.e., station). Surrogate measurements (e.g., conductivity or DOC) may be required as an effluent exposure indicator if more specific plume markers are not available.

Control locations may be difficult to define when receiving waters are subject to other impacts both upstream and downstream. However, a relatively low exposure location can usually be found with some preliminary sampling effort. It may be necessary to relax control acceptance criteria.

Reference toxicants (positive controls) cannot be applied in field exposure situations. However, they can be used in the laboratory as a concurrent check on responsiveness of the test organisms. Similarly, negative controls can be run in the laboratory as a check on health of test organisms, when adequate field control locations are not available.

When field conditions differ from holding conditions, an acclimation period should be allowed in the field at a low exposure location, to prevent acclimation stress effects from affecting the results. It may also be possible to approximate field conditions in a laboratory acclimation system prior to placement of organisms.

In some biomonitoring applications, particularly when potential point sources of toxicity have not been previously identified, dose-response questions are not of primary interest. Rather, the spatial or temporal response pattern is of interest. In these situations, exposure measurement may be less critical, and an LC50 or EC50 result is not desired.

This is similar to the raw sample compliance objective (Section 4.7), except that the study design may include multiple test locations. The response variable is simply a percent mortality (or other endpoint) at each location. A sufficient number of organisms should be placed at each location to give acceptable resolution in the pattern analysis (i.e., confidence in the mortality response at each point).

The need for replication of in situ bioassays should be assessed with regard to the potential for cage effects. In the case of mesh cages of identical size exposed to the same flowing water column, there is probably little need for side-by-side cages. However, in more static water systems, or cages placed in contact with sediment, there may be greater potential for cage-placement or micro-environmental effects, and replicate cages may be desired to permit estimation of such effects.

4.10 Laboratory Evaluation Checklist

A systematic approach is required for evaluation of quality assurance in toxicity laboratories. While specific data quality needs may be project-dependent, a generic list of items to be examined can be constructed. The checklist in Table 4.1 is sufficiently general to meet most project needs. It does not refer to specific tests performed. However, in using the checklist, emphasis can be placed on records and test conditions of particular relevance to the testing services required.

The checklist would be most useful in the hands of a toxicologist performing an on-site inspection. However, it could also be used to assess the completeness of QA program descriptions submitted with proposals. It might also be provided to laboratories as a guide in preparation of QA program descriptions for subsequent evaluation.

Laboratory evaluation checklists can provide a framework for external auditing of toxicity laboratories for purposes of certification (Korndoerfer, 1989) or accreditation. They can also provide a framework for selection of contract laboratories (e.g., NCASI, 1990). However, the intent in this document is mainly to provide guidance and promote consistency in QA program development.

4.11 Residue Toxicity and Bioaccumulation

Traditionally, aquatic toxicity endpoints are expressed in terms of the exposure concentration required to produce a particular effect. However, there may also be a need for corresponding data on the internal chemical residues of the test organism at each effect level. If residues are measured at each exposure concentration, an LD50 or ED50 can be calculated, representing the residue dose required to produce the effect. Residues can differ greatly between tissues and must be clearly stated as either whole-body or specific-tissue residues, on a dry-weight or a wet-weight basis. Whole-body residues are sometimes corrected for gut contents or lipid-normalized, and any such correction must be stated with the residue and LD50 (ED50) results.

Residue toxicity information is primarily useful for interpretation of residue biomonitoring data, and its ecological significance. Residue biomonitoring is sometimes preferred over water quality monitoring of persistent environmental contaminants, if they bioaccumulate well but are difficult to detect in water due to low solubility or intermittent discharge. Residue biomonitoring also permits estimation of contamination transfer through the food chain and beyond the aquatic ecosystem, by application of appropriate bioaccumulation or biomagnification factors.

Aquatic bioaccumulation factors (or biomagnification factors) are experimentally measured in aquatic toxicity laboratories, although bioaccumulation per se is not a toxic, or even necessarily an adverse effect. They are computed either as the steady-state ratio of organism to water contaminant concentrations, or as the ratio of contaminant uptake and depuration rate constants (kinetic method). Methods and QC considerations are provided by the OECD (1981b). Tests can be performed either in static or flow-through mode. Most of the general considerations of Table 4.1 are also applicable in bioaccumulation studies. In addition, some unique considerations are briefly discussed below.

Bioaccumulation studies should be performed well below the toxic concentration range in order to avoid toxic effects on the process (OECD recommends less than 1/10 of the median lethal concentration, but at least 10 times the detection limit in water). Preliminary studies may be needed to estimate the lethal range, as well as the rate of approach to steady-state.

Both steady-state and kinetic methods require concommitant sampling of test organisms, and water, at multiple time points, for chemical analysis. In the steady-state method these are spaced over the steady-state region, which must be long enough to demonstrate equilibrium in organism and water concentrations. The number of samples must be sufficient to achieve specified precision goals for the bioaccumulation factor. This number can be estimated in advance from preliminary data, or from study data in the early steady-state phase, if chemical analysis is rapid. At the end of a static test, a mass balance should be calculated to confirm that the amount of contaminant taken up by the organism approximately equals the amount lost from the water.

In the kinetic method, samples are spaced over the uptake and depuration phases, and are most concentrated over the region of most rapid change. Non-linear regression models are used to simultaneously estimate the uptake and depuration rate constants. Again, the number of samples needed to achieve specified precision goals for the bioaccumulation factor can be estimated in advance from preliminary data, or from study data in the early uptake phase. Kinetic studies are best performed in flow-through mode, since relatively constant water concentrations are required through the uptake and depuration phases.

Bioaccumulation study designs should include two dose levels, in addition to a control, to verify that there are no dose effects on the bioaccumulation process in the concentration range of the study. Static tests should also include an unoccupied exposure vessel at each dose level, in order to determine whether any water concentration changes are due to physical-chemical factors (volatilization, degradation, adsorption on tank walls) rather than biological uptake.

The QA/QC associated with residue sampling and analysis is not unique to bioaccumulation studies and has been described elsewhere (EPA, 1974; FDA, 1975). Replicate sample collections and split samples can be used to estimate sampling method and analytical errors, and analytical performance criteria should be specified. However, experimental design (number and spacing of samples) often limits the total uncertainty associated with a bioaccumulation factor (confidence intervals of ±20% or less are typically achievable). Organisms for analysis should be randomly selected from exposure vessels, and water collected from a central location away from vessel walls. Water should be kept as free as possible of introduced particulates (e.g., waste food and feces).

For analysis of highly adsorptive contaminants both filtered (centrifuged) and total samples should be analyzed. Bioaccumulation factors may be reported relative to either, with full definition of their tissue, weight-basis, and dissolved or total water concentration basis.

TABLE 4.1: QUALITY ASSURANCE CHECKLIST FOR TOXICITY LABORATORIES
(check or enter footnote number to describe deficiencies)

Laboratory Management

_____ Standard Operating Procedures available (both QA and test methods).
_____ Quality Management Plan for laboratory (may be part of SOP).
_____ Responsibilities, authorities, qualifications defined for each position.
_____ Quality Assurance Officer with adequate authority.

Quality Assurance Functions Performed

_____ Review of QC sample results (permanent record available).
_____ Performance assessment review (blind samples).
_____ Knowledge testing and training sessions.
_____ Review of test sample data (adequate checks performed).
_____ Report approval mechanism (signatures).
_____ Quality Assurance reports (corrective actions indicated).
_____ Interlaboratory and accreditation studies (record of participation, performance).

Laboratory Operations

_____ Instrument calibration and maintenance (records, methods available).
_____ Animal care and acclimation (records, methods available, tank materials).
_____ Reagents and working solutions (labels, acquisition, preparation, expiry dates).
_____ Glassware cleaning adequate.
_____ Separation of areas (receiving, storage, preparation, testing, lab floor plan).
_____ Temperature, light, ventilation control (records, adequate devices).
_____ Dilution water supply (source, treatment, chemical check records).
_____ Food supply (source, chemical check records).
_____ Sample information system (log-in, labelling methods, holding times recorded).
_____ Safety procedures adequate.

TABLE 4.1: QUALITY ASSURANCE CHECKLIST FOR TOXICITY LABORATORIES
(check or enter footnote number to describe deficiencies)

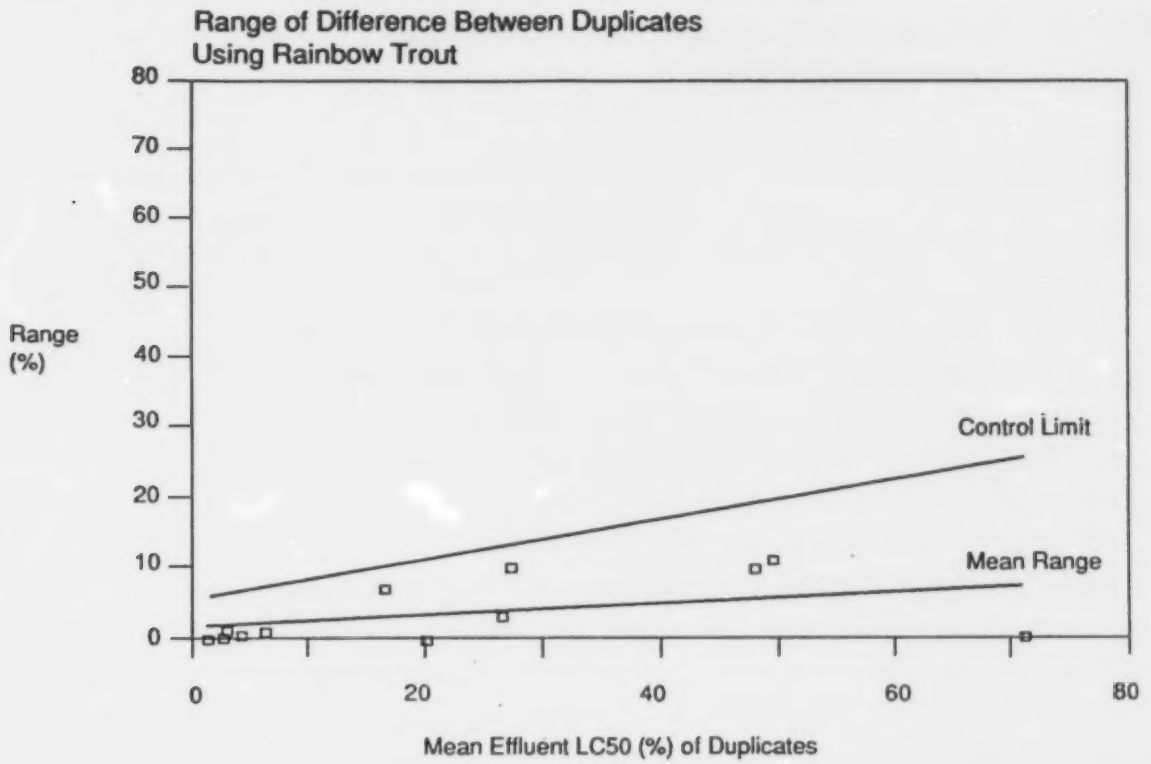
Testing, Quality Control and Reporting

_____ Test record sheets (concentrations, daily responses, test conditions).
_____ Dose selection guidelines (range tests, concentration intervals).
_____ Dose verification (frequency, methods for mixtures).
_____ Dilution water controls (use of data, test acceptance criteria).
_____ Test replication (frequency, range control limits).
_____ Reference toxicant tests (type, frequency, control limits).
_____ Control charts or tables (limit determination, out-of-control actions).
_____ LC50 and other endpoint calculations (method selection guidelines).
_____ Data reporting (confidence limits, methods, test conditions, QC data).

Deficiency Descriptions

Footnote No.	Problem

FIGURE 4.1
Control Charting of Duplicate Toxicity Test Data



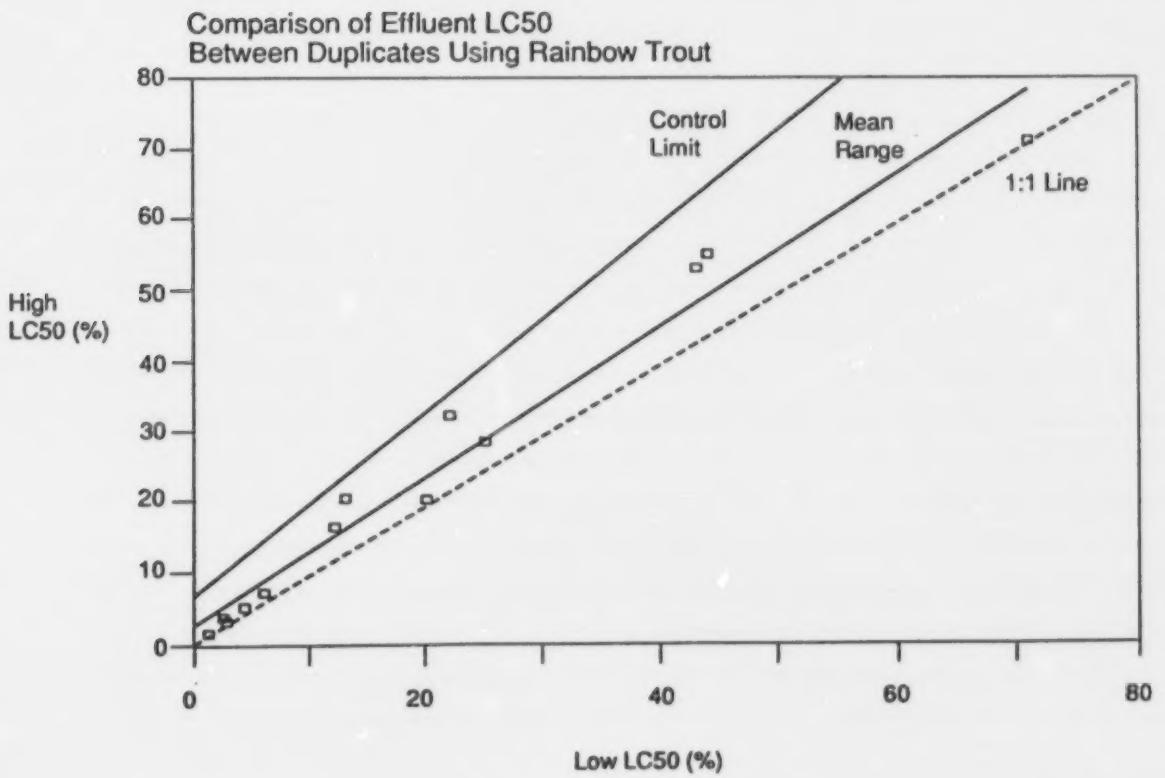
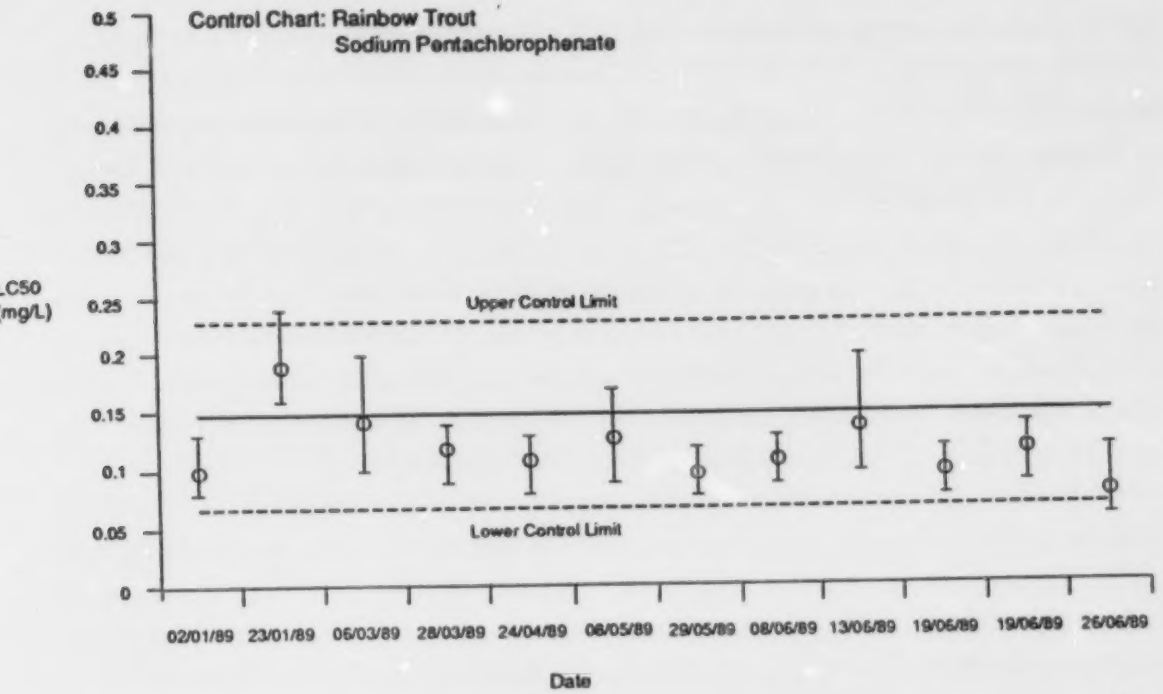
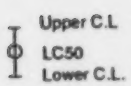
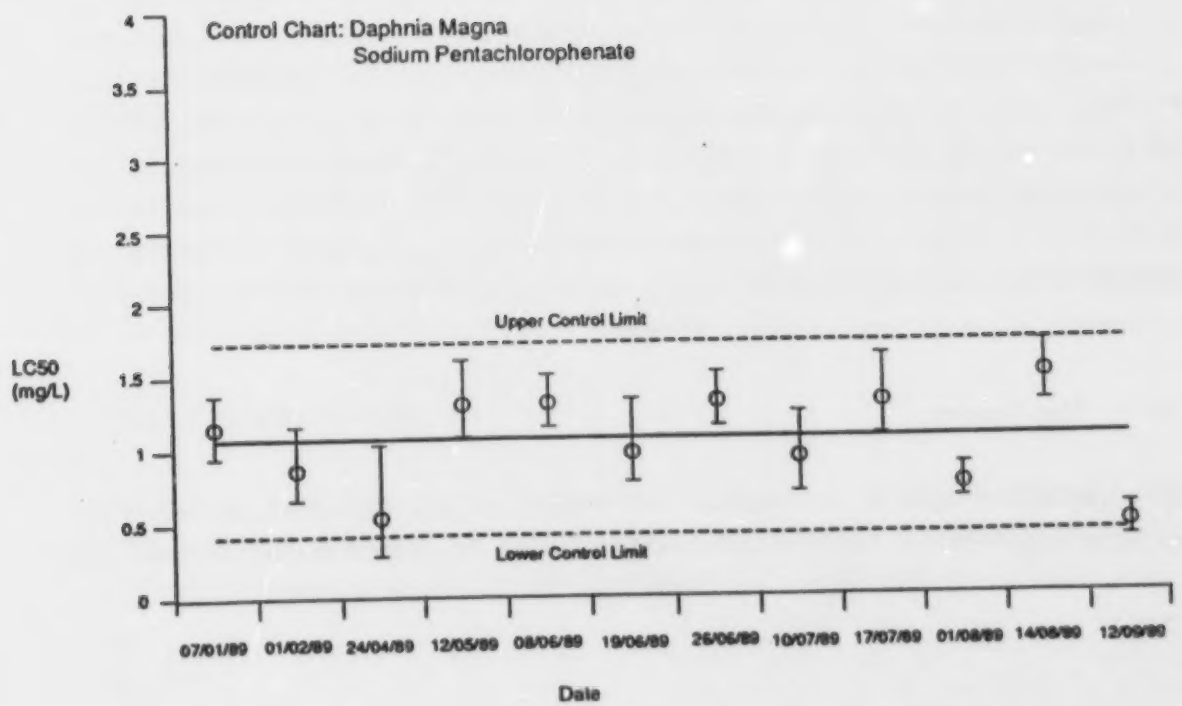


FIGURE 4.2
Average Control Charts for Reference Toxicants







5.0 MICROBIOLOGICAL QA

Microbiological analysis is one type of biological enumeration by which a sample may be structurally characterized. It is similar to chemical characterization in that a sample constituent is being quantified; however, unlike a chemical constituent, a microbial constituent can grow and reproduce within the time-frame of analysis. It must remain alive and must be able to grow on a culture medium to be identified by conventional microbiological methods. These unique aspects of microbiological testing require special precautions to maintain sample integrity and prevent biological contamination. This section focuses on those precautions; however, the general elements of quality assurance outlined in Section 3.0 also apply in the microbiological laboratory.

Specific QA/QC approaches for microbiology are outlined by Prior et al. (1973), Ellis (1976), Bordner (1985) and APHA (1987). The following sections outline recommendations from these sources, with emphasis on environmental rather than chemical microbiology. Prior et al. (1973) provide a laboratory evaluation checklist which is sufficiently broad to encompass both environmental and clinical laboratories.

5.1 Sample Collection and Storage

Sample containers for microbial analysis must be pre-sterilized and aseptic techniques must be followed throughout the sample collection. Samples are generally not preserved, but should be refrigerated below 10°C or placed on ice immediately following collection to minimize microbial growth. Analysis within six hours is recommended by the U.S. EPA and ASTM for environmental samples (Bordner, 1985). Drinking water samples, which are chlorinated, may be held for longer periods, up to 30 hours (Bordner and Winter, 1978; Bordner, 1985). The Ontario Ministry of the Environment recommends analysis of all samples within 24 hours (MOE, 1985b) and quenching of chlorinated samples with sodium thiosulphate. Environment Canada (1983) recommends analysis of all samples as soon as possible within 30 hours.

5.2 Sterilization

As a precaution against contamination of samples or test apparatus, all laboratory surfaces (e.g., benches, floors, walls) should be frequently cleaned and disinfected. This

requires a suitable surface material. Traffic through the laboratory should be limited. A separate laboratory area (autoclave area) should be devoted to sterilization of media and other test apparatus, in addition to the separate areas for washing, (media) preparation and testing that are typical of other laboratories. Standard Operating Procedures should specify times and temperatures for autoclaving media and other materials (e.g., Table A5.1, Appendix 1).

5.3 Methods and Records

Standard test methods for microbiological testing are provided by the U.S. EPA (Bordner and Winter, 1978) and the American Public Health Association (APHA, 1985, 1987). These or equivalent validated methods should be referenced in Standard Operating Procedure documents, and in microbiological data reports.

General methods for testing and servicing of membrane filters and absorbent pads (e.g., Table A5.2), autoclaves and other critical equipment (e.g., Table A5.3) should also be included in Standard Operating Procedure documents, and records of these activities should be maintained in log books. Methods for media preparation should include holding times and conditions, as well as shelf lives for commercial media supplies (e.g., Table A5.4). Detailed records of media preparation should include technicians' names, dates, media lot numbers, dry weights, prepared volumes, sterilization times and temperatures, and pH check data.

Tap or distilled water used in microbiological procedures should be routinely tested for microbial contamination, conductivity, pH and potential toxicants (e.g., Table A5.5). Procedures and frequencies for these tests should be included in Standard Operating Procedure documents, and records of test results should be maintained in log books. Glassware from each wash lot should be spot-checked for beading and pH to guard against media or detergent residues, records of these tests should be maintained, and lots re-washed if residues are found.

5.4 Quality Control

In addition to the general monitoring of water, media, glassware and equipment, specific QC tests should accompany routine tests of microbiological samples with a specified minimum frequency, as per Standard Operating Procedures and/or QA plans.

Sterility Controls

Colony counts should be obtained for test apparatus unexposed to the sample, including media, water, glassware and filters (Membrane Filter tests). These are equivalent to blank samples in chemical analysis, or dilution water controls in toxicity testing, and should be part of every test performed. Acceptance criteria should be defined, and sample data should not be control-corrected.

Duplicate Tests

Bordner (1985) recommends that at least 10% of samples, or at least one sample per batch, should be tested in duplicate. Acceptance criteria for the duplicate range should be defined, along with corrective actions to be taken when these are exceeded.

Control Cultures

Positive and negative cultures should be innoculated into the media periodically to ascertain that the media is performing correctly with respect to survival of the organism being enumerated, and non-survival of other organisms (e.g., Table A5.6). At a minimum, each new media lot should be tested once in this manner. More frequent tests are desirable. These tests are similar to accuracy tests with Standard Reference Materials in chemical analysis, or reference toxicant tests in toxicity testing, but are qualitative rather than quantitative.

Verification Tests

In membrane filter (MF) tests, at least ten colonies monthly, or ten colonies per study, whichever is greater, should be verified biochemically by growth on selective media to confirm their taxonomic identity. In most-probable-number (MPN) tests, at least 10% of positive samples should be verified by processing through the completed test procedures of Bordner and Winter (1978) or APHA (1985).

5.5 Interlaboratory Comparison

Interlaboratory comparison is a desirable part of any quality control program. An example of a formal interlaboratory program which includes microbiological parameters

is the EPA Certification Program for Laboratories Analyzing Drinking Water. The program includes periodic performance evaluation using reference samples unknown to the laboratory, as well as collaborative method studies designed to establish interlaboratory precision of selected methods.

5.6 Viral Testing

Recent studies of treated drinking water and raw water supplies indicate that bacteriophages, and specifically coliphages, are often present in water that is free of total and fecal coliform bacteria (Dutka et al., 1987). The ability of virus to survive conventional water treatments, and to persist for long periods in the absence of host bacteria, suggests that conventional bacteriological parameters may not be adequate indicators of water potability. Coliphage tests (e.g., ASTM, 1982; APHA, 1985) have been shown to be a good indicator of fecal contamination and can be completed more rapidly than coliform tests; therefore, coliphage tests have been recommended for inclusion in microbiological water testing programs (Dutka et al., 1989; Palmateer et al., 1989).

While test protocols are still tentative, the QA/QC approaches described in preceding sections are generally applicable in coliphage testing. Coliphage are cultured in an E. coli host culture in modified tryptic soy agar, and counted as clear plaques in the bacterial lawn. Sterility controls, duplicate tests, and positive and negative culture controls should be frequently performed. The American Type Culture Collection will provide viral cultures. The host culture should be periodically checked for host resistance (challenge old and new host cultures with new viral culture) and for host contamination by virus (spontaneous plaque formation). Verification tests for specific viruses include electron microscopic examination (shape and size), or nucleic acid extraction/amplification and probing with specific monoclonal antibody or DNA/RNA probes.

5.7 Microbial Toxicity Testing

A variety of tests for toxic constituents in environmental samples use bacteria as test organisms (e.g., Dutka et al., 1983; Reteuna et al., 1989). For example, the Microtox assay (deZwart and Sloof, 1983) is based on inhibition of photoluminescence in Photobacterium phosphoreum, and the toxi-chromotest (Fish et al., 1987) is based on

inhibition of a constitutive enzyme system in E. coli. Both tests have method versions for specific detection of genotoxicity. The Ames assay (Ames et al., 1975) is widely used for detection of genotoxicity only.

The U.S. EPA (1978) discusses QA/QC approaches for many microbial toxicity and genotoxicity assays. In addition, OECD (1981b) methods include basic elements of quality control. Generally, the QA principles which apply to toxicity testing (e.g., Section 4.0, Table 4.1, replicates, positive and negative controls) and the specific microbiological precautions of Section 5.0 (e.g., sterility checks, culture controls, etc.) are both applicable in microbial toxicity testing. In genotoxicity testing, it is also important to determine that a negative result is not due to cytotoxicity; genotoxicity testing should always be performed at exposure concentrations well below the cytotoxic range.

6.0 QA IN AQUATIC FIELD STUDIES

A complete biological assessment of pollution impacts on the aquatic environment involves more than toxicity testing of effluent samples. In situ bioassay tests in the receiving water are more definitive for evaluation of short-term effects, while ecological assessment of benthic, plankton and fish communities are more definitive for evaluation of long-term consequences. Benthic communities, because of their sedentary nature, are particularly useful in revealing spatial patterns of biological effect related to point sources of pollution.

A wide variety of sampling and analytical methods are used in ecological community assessments, according to the type of organisms considered, the specific study objectives, and the personal preference of the investigator. With each approach, there are different specific sources of error to be minimized, if possible, and quantified for purposes of data quality evaluation. However, there are common elements and therefore common general approaches to quality assurance, as outlined below. For example, most ecological assessments involve sampling, taxonomic identification and enumeration of organisms. We may be using these data to estimate species biomass, abundance, relative abundance, diversity and/or trophic structure of the community.

The checklist in Table 6.1 includes the same general QA elements from Section 3.0 that were also listed at the top of the aquatic toxicity checklist in Section 4.0. The principles are no less important here, although formal interlaboratory and accreditation studies may be rare. In addition, Table 6.1 includes some new concepts of particular relevance in field ecological assessment. The following paragraphs focus on these field study aspects.

6.1 Preliminary Survey and Design

Development of an optimum sampling design requires preliminary data, to determine, for example, what is an adequate level of sampling effort at each sampling point in time or space. The return of information increases with increasing effort, up to a point of limited return, and this point can be defined. For example, we may choose to collect as many replicate samples as required to limit the increase in total species count at a given station to less than 5% between successive samples (Figure 6.1; Smith, 1974; Stevenson

and Lowe, 1986). Or we may choose to collect as many replicates as needed to achieve a specified level of precision (e.g., relative standard error) for station descriptive statistics (e.g., mean density or diversity). Elliot (1971) describes the simple calculations involved (Appendix 1).

With inadequate effort, we may consistently underestimate the number of species present, or achieve inadequate precision on density estimates. Distribution of the preliminary sampling in space will help to identify any environmental heterogeneities present (i.e., strata), which will then help to optimize design. Sampling and environmental variance can be estimated, and used to pre-determine the appropriate number of samples for the study, or for each stratum, based on project and data quality objectives. In some cases, when sample processing can be completed rapidly in the field (e.g., a fish population estimate, Figure 6.2; Peck et al., 1988), a sequential sampling approach can be used, whereby sampling continues until precision objectives are known to be achieved.

In point-source impact studies, a reference or control area should always be sampled as part of the study design. These areas are located in similar habitat, but beyond the potential impact zone (e.g., upstream). The reference area may simply comprise the outermost stations of the study area (i.e., most distant from the pollution source) with a series of closer stations representing an impact gradient. Gradient studies should be designed to extend beyond the impact zone, and preliminary survey data will be useful in defining this spatial area.

6.2 Training and Performance Evaluation

Crew training and evaluation are particularly important in field studies where written sampling protocols, if they exist, are often lacking in detail and open to some interpretation (Adams et al., 1983). Video recordings can be a useful tool for standardization of field technique, although they are no substitute for writen methods. When several contractors are involved, videos can be required to check and document consistency of technique. Written records should be maintained of crew training and testing sessions, and of remedial actions associated with poor test results. Performance criteria should be defined, including criteria for acceptable error in field measurement and sample processing.

6.3 Sample Collection, Preservation and Storage

Methods will specify sampling equipment appropriate to study objectives, and call for consistency within a project (e.g., USDA, 1983; EPA, 1987). Consistency in equipment is important since different sampling devices have different biases in terms of collection efficiency and type of organism captured. For example, airlift benthic samplers are effective on rocky substrates where dredges will not penetrate, while artificial substrate samplers are highly selective for colonizing organisms. If different types of sampling gear must be used in a study, they should be intercalibrated.

Methods should also specify proper use of equipment, location of stations, sampling season, labelling and preservation of samples, and any constraints on sample storage (e.g., time, temperature). Rapid preservation is important to ensure intact organisms for accurate identification (e.g., 10% neutral buffered formalin for benthic samples). More preservative may be needed in samples containing large amounts of organic debris. Preservation usually occurs prior to sorting organisms from the debris if sorting is required.

6.4 Additional Effort

Additional effort checks are performed in the field, during the routine survey, and serve to confirm the return-for-effort anticipated on the basis of preliminary field work. For example, if a data quality objective is to retrieve 95% of benthic species present at any station, and it has been estimated that three replicates per station are sufficient to achieve this, more replicates can be collected at a few selected stations until the number of species reaches a plateau, and it can be determined whether the first three were sufficient to achieve the 95% objective. The same approach can be applied whether sampling benthos (return vs. number of sediment grabs), plankton (return vs. number of plankton tows) or fish (return vs. number of seine net passes).

Return-for-effort is dependent, in part, on field technique. Thus, additional effort checks can also be used to evaluate field crew performance, in relation to return-for-effort performance criteria.

6.5 Sample Replication

To obtain an estimate of total precision for any species density, relative species density, overall density, or community index, sample collections must be replicated in the field at some or all individual stations. To evaluate spatial patterns, a known confidence is usually desired at every station; thus, all stations are replicated. The number of replicates per station depends on a preliminary estimate of variability, and is chosen to give a desired relative standard error for the mean density or index. A relative error less than or equal to 20% is usually considered acceptable (Elliot, 1971), but this Data Quality Objective should be specified in project QA plans.

6.6 Duplicate Processing and Exchange

Sample processing may involve, as a first step, sorting of organisms from debris, and possibly subsampling of sorted organisms for detailed identification. As with sample collection, written methods for sample processing are required, records should be kept, and errors (i.e., variance) associated with these activities should be estimated (e.g., Kreis, 1986, 1989). Such error estimation necessarily involves some type of duplicate processing, either on split samples, or by re-mixing and reprocessing the same sample (Peck et al., 1988). There are problems associated with each approach (e.g., obtaining an even split or a thorough mix), but these should not preclude at least some effort to characterize processing error, both within and between personnel.

Samples can also be exchanged between laboratories for duplicate processing when several labs are involved in a study. These samples do not have a track record and cannot indicate which laboratory is correct in the event of discrepancy, but they will draw attention to an interlaboratory bias in sorting or taxonomy. Technical workshops may be useful in resolving discrepancies.

Criteria should be established for acceptable duplicate ranges in total organism density, density of individual taxa, or relative species density (e.g., ± 10% of mean). A specified proportion of samples (e.g., 5%) should be processed in this manner, as per QA plans. The data generated by duplicate processing can be used to compare the individual sorters or taxonomists involved in a study, as well as to estimate the overall precision of processing. Individuals who consistently deviate from their peers, causing duplicate

ranges to exceed acceptance criteria, should undergo retraining. As a back-up, the sorted debris should be represerved and kept for later reference. This permits a check on sorting efficiency when discrepancies are observed.

6.7 Sorting Verification

Verification of sorting efficiency is easily performed on a spot check basis if the left-over debris from a sample is retained. Stains can be applied to make any organisms that were missed stand out from the debris (e.g., Eosin B / Biebrich Scarlet). This is particularly useful with benthic samples, which usually contain large quantities of debris. Contract laboratories are much more likely to sort thoroughly when they know that samples will be randomly checked for sorting efficiency. The proportion of samples checked (e.g., 5%) should be specified in QA plans, along with criteria for acceptable sorting performance.

6.8 Taxonomic Verification and Archives

Taxonomic methods will refer to recent published keys, and also describe staining and mounting techniques prior to identification. Keys must be appropriate to the geographic region of study, and should always be cited in data reports. Computerized keys have been developed. Reference collections of organisms should be maintained by each taxonomic lab, with master collections when studies involve several regional laboratories. Photographic iconographs have been used to advantage (Camburn et al., 1984). These collections are useful in confirming identifications and ensuring consistent taxonomy, as well as training personnel.

Verification of taxonomy by recognized experts should be performed on a spot check basis, and on any new specimen entering a reference collection. Museums are sometimes prepared to perform this service when remote areas are included in the study, and new species or distribution records are likely. They will also sometimes serve as archives for selected samples and/or voucher specimens of organisms. However, since considerable effort is involved in long-term maintenance of preserved biological material, it is often necessary to minimize the quantity of material submitted.

Examples of taxonomic keys appropriate to Canadian freshwater invertebrates include Brinkhurst (1986), Clark (1981), Hilsenhoff (1981), Hiltunen and Klemm (1980), Mackie et al. (1980), Merritt and Cummings (1984), Oliver and Rousell (1983), Stimpson and Klemm (1982), Stewart and Stark (1988), Weiderholm (1983), Pennak (1978) and Edwards (1976). For fishes, Scott and Crossman (1973) is widely used.

6.9 Data Handling and Reporting

Ecological community studies usually generate large data matrices, since hundreds of species may be involved, and an observation on each is recorded at every sampling station. Computers and multivariate statistical programs generally assist in synthesis of the data to elucidate spatial and/or temporal patterns of community health (e.g., Hensler, 1982; Ludwig and Reynolds, 1988). Due to the nature and quantity of data, particular care is needed in data handling, to guard against data entry errors or subsequent misuse of data. Double entry systems and transcription checks against the original data records are useful QC techniques. Missing data must be clearly distinguished from a species absence by use of unique non-zero missing value codes, with code definitions built into each file. Read-only files help to ensure data integrity.

6.10 Rapid Bioassessment

Rapid bioassessment methods (e.g., Bode, 1988; Plafkin et al., 1989) are designed for relatively inexpensive generation of a reproducible index of ecosystem health, rather than comprehensive ecosystem description. In these studies, a fixed (and limited) level of effort is desired, rather than a level of effort sufficient to meet pre-defined precision objectives for ecosystem descriptors (e.g., mean ± X%) or information return objectives (e.g., % of maximum number of species). Nevertheless, it is desirable that rapid bioassessment indices be 'reproducible' within certain limits. This reproducibility is usually evaluated in developing the method and its fixed level of effort. Moreover, it is also desirable to document reproducibility (precision) and completeness of information return, on a spot check basis, as part of any rapid bioassessment survey.

An example is the kick-sample method of benthic invertebrate sampling in streams (Bode, 1988). Briefly, it involves kicking the stream substrate for a fixed time period over a fixed-distance downstream transect, and capturing disturbed organisms and debris

in a downstream net of fixed dimensions. Random subsamples of the collected material are examined until a fixed number of organisms are counted and identified. Various simple community indices are computed from the number of species found, and the relative abundances and pollution tolerances of dominant species. For each of these indices, precision measures can be obtained from replicate field samples (e.g., opposite streambanks) and from duplicate processing of selected samples in the laboratory. Percent return of species can be estimated by increased effort at selected stations (e.g., longer time, larger nets) to obtain a maximum species count. Sorting verification can be performed on the debris left over from subsample processing.

Most of the principles and QA approaches described in preceding sections are equally applicable to rapid bioassessment, with a few special considerations. For example, absolute density is not a relevant parameter for precision or performance assessment, since the methods specify a fixed number of organisms examined. Similarly, sampling design does not include level of effort selection (e.g., number of replicates needed achieve a precision objective). Otherwise, the QA checklist in Table 6.1 is sufficiently general to encompass most field biomonitoring programs.

6.11 Macrophyte Surveys

In aquatic macrophyte surveys, the focus of interest tends to be on surface area cover, or vegetation biomass, rather than organism (stem) density. Sampling may involve plots (quadrats) or plotless (distance) methods. In plot sampling, a prescribed area is analyzed at each sampling station (e.g., for % cover, biomass/m^2, density/m^2 of each species). In plotless sampling, cover is estimated from the line-intercepts of species zones on a linear transect, and density from distances between plants at randomly located sampling points. Mueller-Dombois and Ellenberg (1974) provide detailed methods.

Precision estimates are based on replication of either plots or transects within each zone or stratum of the study area to be described. As in any other aquatic field study, the level of effort can be designed, based on preliminary data, to achieve a defined goal for descriptor precision or species return. In vegetation sampling, both plot size (line transect length) and number of plots or line transects are aspects of effort that can be adjusted. Most of the measurements in macrophyte surveys (e.g., cover, density) are performed in the field, and some (e.g., cover) can be rather subjective. Variation

between sampling crews can be assessed and controlled by crew exchange (several crews analyze the same site) when sample exchange is not feasible. Crew training and testing programs should include performance criteria for acceptable discrepancy between crews.

Laboratory processing of samples is usually limited to biomass measurement by species, and confirmation of species identity. Samples are dried to constant weight in drying ovens at low temperature to prevent scorching. Split samples can be utilized to estimate the analytical variance associated with species sorting and biomass measurement. Reprocessing is feasible only as far as wet-weight determination.

Notwithstanding these unique aspects of vegetation sampling, the QA approaches previously described are generally applicable in macrophyte surveys, and the checklist in Table 6.1 can be used in this context.

TABLE 6.1: QUALITY ASSURANCE CHECKLIST FOR AQUATIC FIELD STUDIES
(check or enter footnote number to describe deficiencies)

Quality Management

_____ Standard Operating Procedures available (both QA and test methods).
_____ Quality Management Plan for institution (may be part of SOP).
_____ Responsibilities, authorities, qualifications defined for each position.
_____ Quality Assurance Officer with adequate authority.

Quality Assurance Functions Performed

_____ Review of QC sample results (permanent record available).
_____ Performance assessment review (blind samples).
_____ Knowledge testing and training sessions.
_____ Review of test sample data (adequate checks performed).
_____ Report approval mechanism (signatures).
_____ Quality Assurance reports (corrective actions indicated).
_____ Interlaboratory and accreditation studies (record of participation).

Field Operations

_____ Preliminary sampling (return for effort, environmental strata, variability).
_____ Sampling design (based on objectives, adequate replication).
_____ Instrument calibration and maintenance (records, methods available).
_____ Crew training and evaluation (video records and comparisons).
_____ Sampling equipment (appropriate, consistent, cleaning and maintenance records).
_____ Collection, preservation and storage (methods, adequate labels, records).
_____ Additional effort checks (confirm little gain of information).

TABLE 6.1: QUALITY ASSURANCE CHECKLIST FOR AQUATIC FIELD STUDIES
(check or enter footnote number to describe deficiencies)

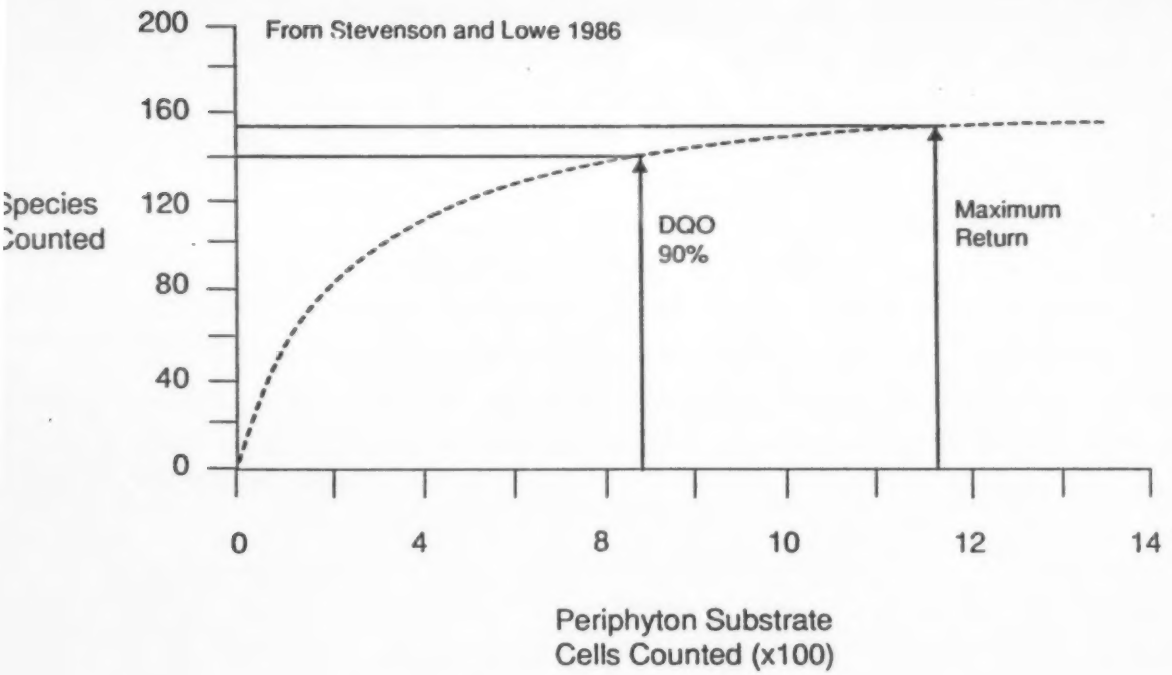
Sample Processing and Reporting

_____ Sorting and subsampling (methods, adequate records, error estimates).
_____ Sorting verification (confirm few organisms overlooked).
_____ Taxonomy and enumeration (appropriate keys, records, reference collection).
_____ Duplicate processing (within and between personnel, error estimates).
_____ Exchange samples (between laboratories, comparability assessed).
_____ Taxonomic verification (by recognized experts).
_____ Archives (for samples, voucher specimens).
_____ Data handling and reporting (entry checks, missing value codes).

Deficiency Descriptions

Footnote No.	Problem

FIGURE 6.1
Designing Effort to a Species Inventory Objective



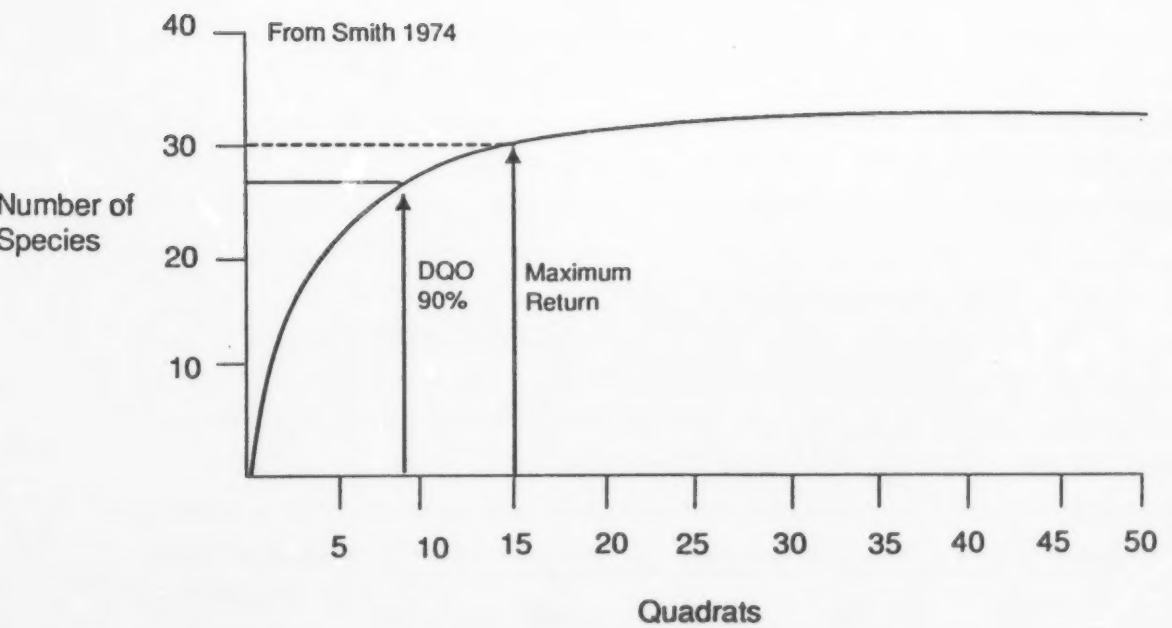
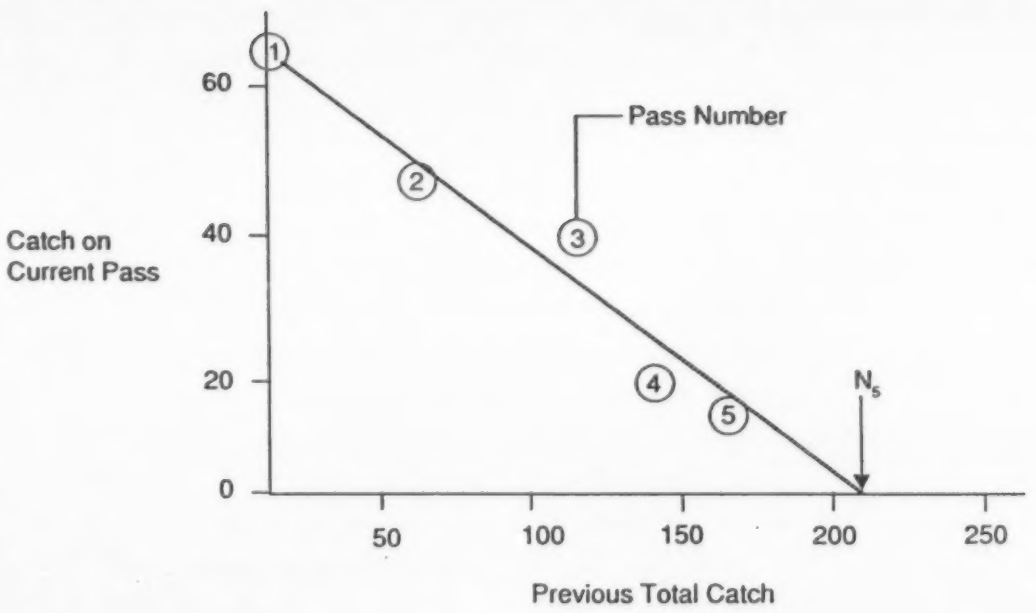
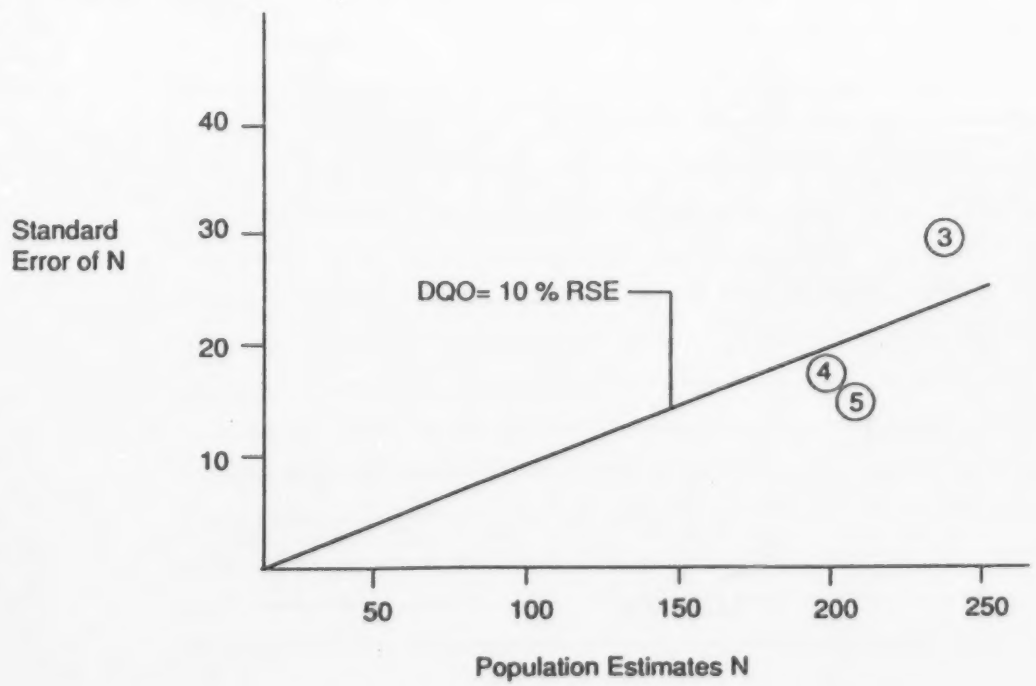


FIGURE 6.2
Sequential Sampling to Meet a Precision Objective





7.0 QA IN PALEOLIMNOLOGY J.P. Smol, B.F. Cumming, and J.C. Kingston

Effective ecosystem management requires (Ford, 1988):

1. knowledge of baseline conditions of natural variability;
2. identification of the point in time that we can confidently say the system began to change; and
3. description of the range of probable trajectories.

Because historical monitoring data are rarely available, paleolimnological studies are being used increasingly for environmental assessments of lake systems. Traditionally, most of these studies have focused on environmental problems such as eutrophication and acidification, but paleolimnological techniques are now being transferred to many new research areas, such as studies of global climate change. These studies often encompass broad geographical areas and involve large project teams, and QA/QC considerations form an essential component.

Paleolimnological protocols and associated QA guidelines have been developed for two large projects which investigated effects of acid rain on aquatic resources: the Paleoecological Investigation of Recent Lake Acidification (PIRLA) project in the United States, and the Surface Water Acidification Programme (SWAP) in Great Britain and Scandinavia (Battarbee and Renberg, 1985; Charles and Whitehead, 1986a,b; Kreis, 1986; Kingston et al., 1988; Cumming et al., 1989; Smol et al., 1989; Charles and Smol, 1990; Munro et al., 1990). While these are generally consistent with the QA guidelines of Section 3.0, this section emphasizes some key aspects of method selection, documentation and quality assurance that are specific to paleolimnological studies, based on PIRLA and SWAP experience.

7.1 Responsibilities and Authorities

A hierarchical structure of clearly defined reporting channels and regular reporting schedules is recommended when many organizations and individuals are involved in the project. Frequent planning, data summary and taxonomy sessions are critical to effective large project coordination. QA plans and QA reports should be available to all participants.

7.2 Sampling Procedures

Most paleolimnological methods have been thoroughly evaluated and standardized (Berglund, 1986; Charles and Whitehead, 1986b). Nevertheless, improvements are actively sought and tested (e.g., Glew, 1988; Birks et al., 1990). Written documentation of methods is essential to consistent performance. Methodologies from the PIRLA project include lake selection, watershed and climatic histories, coring, dating, core subsampling and archiving, and various taxonomic and chemical methods manuals. Guidelines on coring method selection and field documentation are summarized below.

Coring and Sectioning

Coring and sectioning methods will depend on the objectives of the project, and methods will vary depending on the time frame of interest (i.e. the depth of penetration required), sediment characteristics, and the depth of the lake from which the sediments are to be removed (McKean and Nordin, 1986). Appropriate methods differ for short-term (i.e., last 150 years) and long-term studies (i.e., the last 12,000 years).

Short-term Coring Methods (ca. last 150 years)

Most enviromental assessments in North America are interested in time scales of one or two centuries, and so "short" sediment cores (i.e. less than one metre in length) are required. If one is interested in recent stratigraphic changes, it is recommended that either cable-operated piston cores, gravity cores, or freeze samples be taken (Aaby and Digerfeldt, 1986). Gravity coring is a simple, efficient, and widely used method of obtaining sediment samples; however, this technique requires considerable expertise to obtain satisfactory cores. In PIRLA II, a modified K-B gravity corer (Brinkhurst et al., 1969, Brinkhurst, 1974; Glew, 1989) was used (Smol et al., 1989).

A minimum of two cores should be collected at each lake to ensure that sufficient sediment is available in the event that the primary core cannot be analyzed. Back-up cores can also be used for additional study of within-lake variability. More information on these procedures can be found in Charles and Whitehead (1986b) and Smol et al. (1989).

Significant disturbance in a sediment stratigraphy may occur if cores are transported prior to sectioning. To avoid these problems sediment cores should be sectioned at the lake shore. If a major objective of this research is to detect very recent trends (i.e. the last 20 years), an extrusion system which allows close-interval sectioning of sediment cores with negligible vertical contamination should be used. The Glew (1988) sectioner has been used very successfully in a large number of detailed studies.

The usefulness of freeze samplers and tape-peel slide preparation methods in paleolimnological studies has been well documented (e.g., Renberg, 1981, Simola, 1986), especially when a continuous record of past lake changes is required. Finer temporal resolution is potentially possible using Davidson's (1988) modification of Simola's (1986) procedure.

Long-term Coring Methods (ca. last 12,000 yrs.)

A number of coring procedures are available for obtaining a continuous record of sediment. The Livingstone (1955) corer is the most frequently used corer in freshwater studies but is difficult to use effectively when water depths exceed 20 m (Mackereth, 1958; Wright et al., 1965). Several modifications have been made to the original Livingstone (1955) design and Aaby and Digerfeldt (1986) recommend Wright's (1967) modification. Continuous corers such as the one designed by Mackereth (1958), Digerfeldt (1978), Gilbert and Glew (1985), or McKean and Nordin (1986) can also be useful for specific applications. Whatever the method chosen, one should be aware of the particular problems and advantages of each. For more information on coring procedures, see Aaby and Digerfeldt (1986).

Field Notes and Organization

Well-organized, accurate, and complete field notes are essential for paleolimnological studies, and an important QA consideration. These notes will help determine which cores will be given priorty for analyses in the lab. In addition they will provide valuable information that will be required in subsequent interpretations.

All field notes should include, at an absolute minimum, the following information:

- The date and time when the lake was cored.
- The personnel of the coring crew, identifying the team leader. The leader must have extensive coring experience.
- Any pertinent information pertaining to weather conditions which could either help or hinder the coring procedure.
- The type of corer used.
- A morphometric map (or sketch) showing the location of the coring site.
- The depth from which the core was obtained.
- The length of the core.
- Detailed description (including photographs) of the core describing the nature of the sediment/water interface, any colour changes in the sediment, laminations (if present), core consistency, and any observations on which the integrity of the core can be judged (e.g. evidence of burrowing invertebrates).
- Detailed notes on the sectioning of the cores. These notes should include the intervals at which the core was sectioned, noting any problems encountered during the sectioning. Additional notes assessing core length and core integrity should also be logged.
- Sectioning notes should also include what type of bags or vials were used to store the sediment and how they were labelled. Labels must be written twice (using permanent, water-proof ink) including the lake name, the date when the core was taken, the sectioned interval, and the core number.
- While at the lake, additional notes on vegetation should be made, as well as a description of the lake basin. Evidence of any recent changes in lake level (e.g., drowned trees, beaver dams etc.) should also be noted. Photographs should also be taken, and logged in the field notes.
- Any additional samples that were collected should be noted (e.g. phytoplankton, periphyton, zooplankton, and macrophyte samples, etc.).

Before leaving the coring site, the leader must double check that all information was correctly logged. Upon returning to the laboratory, all notes must be typed and collated. The original field notes are photocopied. One copy should be archived and available to others if required (e.g., Cumming et al., 1989).

7.3 Analytical Procedures

Analytical procedures, including chemical and taxonomic/enumeration methods, also require complete written documentation. Methodological references from the PIRLA program in the important areas of sediment dating and plankton enumeration are briefly described below.

Plankton Analysis

While various taxonomic groups of plankton may be of interest, diatoms and chrysophytes have traditionally formed the mainstay of paleoenvironmental assessments because:

- o their siliceous cell wall components that are abundantly preserved in lake sediments are taxonomically specific;
- o taxa within these two groups are ecologically diverse responding to a large number of environmental variables; and
- o the environmental tolerances and optima of these taxa can be quantified to a high degree of certainty using robust statistical analyses.

Diatom and chrysophyte analyses follow the basic protocols outlined in Kingston (1986) and Smol (1986), respectively. Digestion procedures should be documented and validated (e.g., Smol, 1983). Enumeration of plankton remains must be done on high quality research microscopes (e.g., 100x oil objectives, NA 1.3-1.4). Scanning and transmission electron microscopy should be used when required to confirm taxonomic identifications.

Sediment Dating Analysis

There are a wide variety of methods used to date lake sediment (Olsson, 1986; Thompson, 1986; Saarnisto, 1986; Berglund, 1986). Whenever possible, several independent methods should be used; the choice of dating techniques will depend on the objectives of the study. For each particular technique, appropriate QA/QC protocols should be observed, and a separate QA/QC document should be prepared by the dating facility.

7.4 Quality Control

Sampling Quality Control

Coring QC

Coring QC is designed to evaluate and separate components of variability due to differences in sediment stratigraphy at different locations within a lake, and differences due to the coring process itself. The design of a QA/QC program will depend on the size and objectives of a study. For example, Charles and Smol (1990) addressed problems related to the variability of surface sediments from different locations in a lake, and the reliablity of a single sediment core from the deepest basin of an acidic lake. Davis et al. (1990) compared stratigraphic interpretations from triplicate cores from acidic lakes in northern New England. Anderson (1989) undertook a study of multiple sediment coring in a recently eutrophied lake system. Kreis (1986, 1989) specifically addresses potential sources of variaolity encountered during coring and sectioning for the PIRLA I project.

Workers generally show that single sediment cores from the deepest basin are adequate for overall lake measures of diatom community ecology (Kreis, 1989; Charles and Smol, 1990; Cumming et al., 1990). Cores should always be checked on retrieval for any evidence of disturbed lamination.

Preventative Maintenance

Coring equipment should be cleaned and inspected before and after each use. In addition, identical back-up corers and extruders should always be transported to the study area.

Analytical Quality Control

Plankton Enumeration

Plankton QC should consist of, at a minimum, the following:

1. preparation of a compendium of photographic plates of taxa encountered (taxonomic consistency);

2. replicate analyses of the same microscope slide by one counter (intra-slide variability);
3. replicate analyses of the same slide enumerated by different counters (inter-counter variability);
4. replicate analyses of different microscope slides made from the same sediment sample (inter-slide variability); and
5. replicate analyses of microscope slides prepared from different aliquots from the same sediment increment (intra-core variability).

These QC activities generally consume approximately 10-25 % of the overall research effort.

Photographic plates (e.g., Camburn et al., 1984-1986) are important because the biggest source of inaccuracy in this type of research comes from inconsistent or inaccurate taxonomic identifications. Although preparation of photographic plates is time-consuming, it provides a medium for taxonomists to maintain consistency in taxonomy, and allows future workers to more readily compare their data to past analyses. In PIRLA, mandatory taxonomic workshops were held every six months. Similarly, taxonomic consistency was also a primary goal in SWAP (Munro et al., 1990). Floras resulting from these studies should be published when possible, and shorter taxonomic references are encouraged.

Intra-slide variability can be assessed by replicate counts of the same slide by a single counter. This is important because, in a typical count, only a limited number of fields are enumerated. Therefore variability in slide consistency must be assessed (Kreis, 1986; 1989; Charles and Smol, 1990).

Inter-counter variability is important to assess because of possible counter biases due to taxonomic inconsistencies, etc. It is crucial to maintain consistency when many counters are involved in a single project. This involves multiple counts by different counters on single slides from a representative set of lakes, followed by evaluation of the results and appropriate changes in taxonomic identifications (Kreis, 1986, 1989; Cumming et al. in Sullivan, 1990).

Inter-slide variability is assessed with repeated slide preparations from the same sedimentary aliquot. Intra-core variability is the variability associated with slide preparation from different sediment samples of the same stratigraphic interval (Kreis, 1986, 1989; Charles and Smol, 1990).

Although there are several potential sources of variability in all biomonitoring work, including paleolimnology, research during the PIRLA and SWAP projects has repeatedly demonstrated that this variablity is relatively small when appropriate QA/QC protocols are followed. Such an extensive analysis of potential sources of variablity in biomonitoring has only been undertaken in paleolimnologcial analyses of diatom and chrysophytes. Approaches used in paleolimnology have been independently validated by comparisons of reconstructed variables with known changes in water chemistry from long-term monitoring programs (Dixit et al., 1989).

Sediment Dating (^{210}Pb Analysis)

An adequate QA program should accompany the radionuclide analysis. For example, the Chalk River Nuclear Laboratory performs the following QA tests on ^{210}Pb analysis:

- Cross-reference samples of yield tracer solution (Po-209) with analysis by another laboratory (Freshwater Institute, Winnipeg).
- Verify the activity of the ^{209}Po solution against an unknown ^{208}Po solution National Bureau of Standards and against known ^{210}Po solutions.
- Measure the activity of ^{210}Pb on core samples supplied by the Energy, Mines and Resources.
- Run blanks, yield tracer solutions and reagents through the analytical procedure in order to detect any contamination with ^{210}Po or ^{210}Pb in any analyses.
- The background on the spectrometers is monitored at approximately two-week intervals. The background activity is generally less than 1 % of the counts.

Radioanalytical dating can also be independently corroborated by pollen dating on selected cores (Hart and McKee, 1985; McKee et al., 1987). McAndrews et al. (1973) provide pollen keys for the Great Lakes region, and Stockmarr (1970) describes a spore

spiking technique which permits recovery correction to obtain absolute pollen densities. The weed pollen horizon (1830-40) is a useful marker in southern Ontario.

Preventative Maintenance and Frequency

Research microscopes should be routinely maintained before and after each use; professional maintenance in response to serious difficulties (for example cleaning an objective internally) is done by the product vendors on an "as needed" basis, perhaps every five years. Likewise, the ^{210}Pb analysis equipment used must be continuously maintained; problems are recognized if and when calibrations of standard samples deteriorate, and remedial action is taken before any new samples are analyzed. A logbook of all maintenance and equipment changes should be kept.

Corrective Action

Appropriate corrective actions should be taken if errors in analytical techniques are discovered. These may include, but are not limited to, recounting certain taxa on slides if incorrect taxonomic designations (usually of newly encountered taxa) are discovered, subsequent correction of data files, and repeated statistical analyses based on those corrections. The back-up core can be analysed if problems are discovered with the primary core. As with all actions undertaken in a QA/QC plan, details concerning the problem and corrective action taken should be documented.

Statistical Evaluation

Effort should be made to keep abreast of ecologically valid, up to date methods of statistical inference. A statistician should be consulted, to review experimental design and assist with data analysis. A number of multiple regression methods utilizing categorical grouping of species are now considered outdated and should not be used (see Birks et al., 1990). Ordination techniques are useful for interpretation of multi-species ecological data (Jongman et al., 1987; Ter Braak, 1987). In addition, there are now a number of computer intensive methods (e.g., Monte Carlo simulation, bootstrapping, etc.) which can be utilized to develop reliable error estimates for paleolimnological inferences (Birks et al., 1990).

Personnel and Slide Exchange

Comparisons between colleagues and institutions involved in a large program are facilitated by attendance at meetings and workshops where microscope slides are examined, and exchanged. Investigators involved in this type of work must have a firm commitment to attending and contributing to these workshops. For example, during PIRLA I, mandatory taxonomic workshops were held every six months.

7.5 Documentation Procedures

Data Reporting and Handling

QC and routine data will be obtained in three main categories: (1) plankton data; (2) dating information; and (3) water chemistry and other lake and watershed information. Data entering the data base from the original count sheets must be checked for correct entry by data input personnel. These counts should then be sent to the original counter in order to ensure accurate input. The creation of a powerful and flexible data base management system requires a great deal of time and effort but is very useful for large cooperative projects. PIRLA uses the Scientific Information Retrieval system (SIR) which provides output of system files for direct statistical analysis by SAS, SPSS, BMDP, and other packages (Charles et al., 1989).

Sample Custody

Sediment cores should be subsampled at intervals and subsamples archived according to documented procedures (e.g., Charles and Whitehead, 1986b). Sediment subsamples should be archived in a cold room, for future analysis. The coldroom temperature should be constantly monitored by an alarm system. Sediment samples should be stored in appropriate containers at 4°C along with wet paper towels to retard drying.

Complete records should be maintained to identify and track sediment samples. These include field and coring data forms, subsampling and custody forms. An example of the form used in the Queen's University paleolimnology lab is included as Figure 7.1.

At the end of any project, one complete set of slides, counts and ecological data should be archived at a national museum or recognized collection (e.g. the National Museum of Natural Sciences, the Academy of Natural Sciences of Philadelphia), thereby making them available for future analyses by the scientific community. The final repository of the samples and slides should be stated in the publication and reports resulting from the studies.

FIGURE 7.1 :
CUSTODY RECORD - LAKE SEDIMENT SAMPLES
Paleoecological Investigation of Recent Lake Acidification II

This form is to be used whenever sediment is removed from Queen's Univ.

Note: When information is added to this form, return a copy to the originator.

Originator:
John P. Smol
Dept. of Biology
Queen's University
Kingston, Ontario
K7L 3N6
(613) 544-6147

Lake Name - _____
Date cored - _____
Core # - _____
Sent by - _____
Date sent - _____
Received by - _____
Date received - _____
Date returned to Queen's - _____
Sediment Condition (circle one) - (wet, dry, ash)
Container (circle one) - whirlpak, scintillation vial

Intervals sent	Weight of interval (g)	Amount of sediment removed (g)

8.0 REFERENCES

APHA. 1985. Standard Methods for the Examination of Water and Wastewater. 16th Ed. American Public Health Association, Washington, D.C.

APHA. 1987. Standard Methods for the Examination of Water and Wastewater. 17th Ed. American Public Health Association, Washington, DC.

ASTM. 1982. Standard test method for coliphages in water. D4201-82, ASTM Annual Book of Standards. American Society of Testing and Materials, Philadelphia, PA.

ASTM. 1985. Standards on Precision and Bias for Various Applications. Second Edition. American Society for Testing and Materials, Philadelphia.

ASTM. 1986. Manual on Presentation of Data and Control Chart Analysis. Committee E-11 on Statistical Methods. ASTM Special Technical Publication 15D. American Society for Testing and Materials.

ASTM. 1988a. Draft Standard Practice for Generation of Environmental Data. Committee E-47 on Biological Effects and Environmental Fate. Subcommittee E47.05 on Quality Assurance. American Society for Testing and Materials, Philadelphia.

ASTM. 1988b. Standard Practice for Conducting Acute Toxicity Tests on Wastewaters with Daphnia. D-4229-84. In: Annual Book of ASTM Standards, Vol. 11.04. American Society for Testing and Materials, Philadelphia.

ASTM. 1988c. Standard Practice for Conducting Acute Toxicity Tests with Fishes, Macroinvertebrates, and Amphibians. E-729-80. In: Annual Book of ASTM Standards, Vol. 11.04. American Society for Testing and Materials, Philadelphia.

ASTM. 1988d. Standard Practice for Standardized Aquatic Microcosm: Fresh Water (Draft). American Society for Testing and Materials, Philadelphia.

ASTM. 1989a. Guide for Sediment Collection, Storage, Characterization and Manipulation (Draft). American Society for Testing and Materials, Philadelphia.

ASTM. 1989b. Conducting Solid-phase Sediment Toxicity Tests with Freshwater Invertebrates (Draft). American Society for Testing and Materials, Philadelphia.

ASTM. 1989c. Proposed Guide for Conducting Solid-phase 10-day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods (Draft). American Society for Testing and Materials, Philadelphia.

Aaby, B. and G. Digerfeldt. 1986. Sampling techniques for lakes and bogs. In: Berglund, B.E. (ed.) Handbook of Holocene Palaeoecology and Palaeohydrology. pp. 181-194.

Adams, W.J., T.J. Hoogkeem and P.R. Michael. 1986. Aquatic Monitoring: A Rationale for Obtaining and Interpreting Aquatic Ecosystem Chemical Exposure Data. In: Isom, B.G. (Ed.), Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems, STP 894. American Society for Testing and Materials, Philadelphia.

Alberta. 1978. Wastewater Effluent Guidelines for 96-hour Multiple Concentration Static Bioassay Using Rainbow Trout. Alberta Environment, Pollution Control Division.

Ames, B.N., J. McCanna nd E. Yamasaki. 1975. Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome mutagenicity test. Mutation Res. 31: 347-364.

Anderson, N.J. 1989. A whole-basin diatom accumulation rate for a small eutrophic lake in northern Ireland and its palaeoecological implications. Journal of Ecology 77: 926-946.

Atkinson, G. 1985. ODCA Sampling. Project U390 (draft). Environmenal Protection Service, Ottawa.

B.C. MOE. 1982. Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish. File SU-64.1205. British Columbia Ministry of the Environment.

Battarbee, R.W. and I. Renberg. 1985. Royal Society Surface Water Acidification Project (SWAP) palaeolimnology programme. Research paper no. 12 of the Palaeoecology Research Unit (University College London).

BEAK. 1990. Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants (draft). Report to Environment Canada, Conservation and Protection, by Beak Consultants Limited, Brampton, Ontario.

Berglund, B.E. 1986. Handbook of Holocene Palaeoecology and Palaeohydrology. John Wiley and Sons. N.Y. 869 pp.

Birks, H.J.B., J.M. Line, S. Juggins, A.C. Stevenson, and C.J.F. ter Braak. 1990. Diatoms and pH reconstruction. Phil. Trans. Roy. Soc. Lond. B (in press).

Bode, R.W. 1988. Methods for Rapid Biological Assessment of Streams. New York State Department of Environmental Conservation, Albany.

Bordner, R.H. 1985. Quality assurance for microbiological analysis of water. In: J.K. Taylor and T.W. Stanley (Eds.). Assurance for Environmental Measurements. ASTM STP 867. American Society for Testing and Materials, Philadelphia.

Bordner, R.H. and J.A. Winter (Eds.). 1978. Microbiological Methods for Monitoring the Environment, Water and Wastes. EPA-600/8-78-017. Environmental Monitoring and Support Laboratory, Environmental Protection Agency, Cincinnati, OH.

Botts, J.A. et al. 1989. Toxicity Reduction Evaluation Protocol for Municipal Wastewater Treatment Plants. U.S. EPA Risk Reduction Engineering Lab. EPA-600-2-88-062. Cincinatti, OH.

Bouck, G.R. 1984. Physiological responses of fish: problems and progress toward use in environmental monitoring. In: Contaminant Effects on Fisheries, V.W. Cairns et al. (Eds.). John Wiley and Sons, New York.

Brinkhurst, R.O. 1986. Guide to the freshwater aquatic microdrile oligochaetes of North America. Can. Special Pub. Fish. Aquat. Sci. 84: 259. Ottawa, Canada.

Camburn, K.E., J.C. Kingston, and D.F. Charles (eds.). 1984-1986. PIRLA Diatom Iconograph. Contains 53 photographic plates and 1059 figures, plus figure legends. PIRLA Unpublished Report Series, Report number 3. Indiana University, Bloomington, IN.

Charles, D.F. and J.P. Smol. 1990. The PIRLA II project: regional assessment of lake acidification trends. Verh. Internal. Verein. Limnol. 24 (in press).

Charles, D.F., J.P. Smol, A.J. Uutala, P.R. Sweets, and D.R. Whitehead. 1989. The PIRLA database management system. INQUA Newsletter no. 2: 3-6.

Charles, D.F. and D.R. Whitehead. 1986a. The PIRLA project: Paleoecological investigations of recent lake acidification. Hydrobiologia 143: 13-20.

Charles, D.F. and D.R. Whitehead (Eds.). 1986b. Paleoecological Investigation of Recent Lake Acidification (PIRLA): Methods and project description. Electric Power Research Institute, Palo Alto, CA.

Clarke, A.H. 1981. The Freshwater Molluscs of Canada. National Museum of Natural Sciences Canada. 446 p.

Cumming, B.F., A.J. Uutala, S.S. Dixit, J.P. Smol, and J.C. Kingston. 1989. Coring field notes - PIRLA II, Adirondacks. Unpublished report no. 2.

Cumming, B.F., J.P. Smol, J.C. Kingston, S.S. Dixit, A.J. Uutala and D.F. Charles. 1990. Paleolimnological Variability. In: Sullivan, T. 1990. Historical changes in surface water acid-base chemistry in response to acidic deposition. State of Science/Technology Report 11, NAPAP.

Davidson, G.A. 1988. A modified tape-peel technique for preparing permanent qualitative microfossil slides. J. Paleolim. 1: 229-234.

Davis, R.B., D.S. Anderson, M.C. Whiting, J.P. Smol and S.S. Dixit. 1990. Alkalinity and pH of three lakes in northern New England, U.S.A. over the past 300 years. Phil. Trans. Roy. Soc. Lond. B (in press).

Day, K.E., E.D. Ongley, R.P. Scroggins and H.R. Eisenhauer. 1988. Biology in the New Regulatory Framework for Aquatic Protection. Proceedings of the Alliston Workshop, Environment Canada, Ottawa.

deZwart, D. and W. Sloof. 1983. The MICROTOX as an alternative assayin the acute toxicity assessment of water pollutants. Aquatic Toxicity. 4: 129-138.

Digerfeldt, G. 1978. A simple corer for sediment sampling in deep water. Rep. 14. Dep. Quaternary Geol., Univ. Lund, Sweden. 8 pp.

Dixit, S.S., A.S. Dixit, and J.P. Smol. 1989. Lake acidification recovery can be monitored using chrysophycean microfossils. Can. J. Fish. Aquat. Sci. 46: 1309-1312.

Dorn, P.B. and J.H. Rodgers, Jr. 1989. Variability associated with identification of toxics in national pollutant discharge elimination system (NPDES) effluent toxicity tests. Environmental Toxicology and Chemistry 8: 893-902.

Dutka, B.J., N. Nyholm and J. Petersen. 1983. Comparison of several microbial toxicity screening tests. Water Res. 17: 1363-1368.

Dutka, B.J., G.A. Palmateer, S.M. Meissner, E.M. Janzen and M. Sakellaris. 1989. Coliphage and bacteriophage as indicators of groundwater quality in Canada. NWRI Contrib. #89-149. National Water Research Institute, Burlington.

Dutka, B.J., A. El Shaarawi, M.T. Martins and P.S. Sanchez. 1987. North and South American studies on the potential coliphage as a water quality indicator. Water Research 21: 1127-1135.

Dux, J.P. 1986. Handbook of Quality Assurance for the Analytical Chemistry Laboratory. Van Nostrand Reinhold Co., New York.

Ellis, R.J. 1976. Manual of Quality Control Procedures for Microbiological Laboratories, 2nd edition. Center for Disease Control, U.S. Department of Health, Education and Welfare, Atlanta.

EPA. 1974. Section 5, A, (1), Analysis of Human or Animal Adipose Tissue, In: U.S. Environmental Protection Agency. Analysis of Pesticide Residues in Human and Environmental Samples. Thompson, J.F., ed. Research Triangle Park, N.C.

EPA. 1987. Quality Assurance and Field Methods Manual. Illinois Environmental Protection Agency, Division of Water Pollution Control. Chicago.

EPA. 1978. Quality Assurance Guidelines for Biological Testing. Prepared for EPA Environmental Monitoring and Support Lab, Las Vegas, by Tractor Jitco, Inc., Rockville, Maryland. PB-285 369.

EPA. 1979. Interim NPDES Compliance Biomonitoring Inspection Manual. Environmental Protection Agency, Washington. PB81-116808.

EPA. 1983/87. Toxic Substances Control Act (TSCA) Good Laboratory Practice Standards. Environmental Protection Agency. Final Rule. Federal Register 48(230): 53922-53944. Proposed Rule. Federal Register 52(248): 48933-48946.

EPA. 1984. Definition and Procedure for the Determination of the Method Detection Limit - Revision 1.11. Appendix B to Part 136. Federal Register 49(209): 26 October 1984.

EPA. 1985a. Methods for Measuring the Acute toxicity of Effluents to Aquatic Organisms. Third Edition. (W. Peltier and C.I. Weber). EPA Office of Research and Development, Cincinatti.

EPA. 1985b. Methods for Measuring the Chronic toxicity of Effluents to Aquatic Organisms. Third Edition. (W. Horning and C.I. Weber). EPA Office of Research and Development, Cincinatti.

EPS. 1980. Standard Procedure for Testing the Acute Lethality of Liquid Effluents. Environmental Protection Service, Water Pollution Control Directorate. EPS-1-WP-80-1.

Edmunds, G.F., Jr., S.L. Jensen and L. Berner. 1976. Mayflies of North and Central America. University of Minnesota Press, Minneapolis. 330 p.

Elliot, J.M. 1971. Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association, Scientific Publication No. 25.

Environment Canada. 1983. Sampling for Water Quality. Water Quality Branch, Inland Waters Directorate, Ottawa.

Environment Canada. 1989a. Acute Lethality Test Using Rainbow Trout (Salmo gairdneri). Final draft, April 1989.

Environment Canada. 1989b. Acute Lethality Test Using Daphnia sp. Final draft, April 1989.

Environment Canada. 1989c. Acute Lethality Test Using Threespine Stickleback (Gasterosteus aculeatus). Final draft, April 1989.

FAO. 1987. Manual of Methods in Aquatic Environment Research. Part 10 - Short-term Static Bioassays. (D.L. Reish and P.S. Oshida). Food and Agriculture Organization of the United Nations, Rome.

Fava, J.A. et al. 1988. Generalized Methodology for Conducting Industrial Toxicity Reduction Evaluations (TREs). U.S. Environmental Protection Agency, Risk Reduction Engineering Lab. EPA-600-2-88-070, Cincinatti, OH.

FDA. 1975. Revision. Pesticide Analytical Manual, Vol. 1. U.S. Food and Drug Administration, Rockville, MD.

Finney, D.J. 1981. Probit Analysis. Cambridge University Press, Cambridge.

Fish, F., I. Lampert, A. Halachmi, G. Riesenfeld and M. Herzberg. 1987. The SOS chromotest kit: a rapid method for the detection of genotoxicity. Toxicity Assessment: An International Quarterly. 2: 135-147.

Ford, J. 1988. The effects of chemical stress on aquatic species composition and community structure. In: Levin, S.A. et al. (eds.), Ecotoxicology: problems and approaches. Springer-Verlag, N.Y. pp. 99-144.

Giesy, J.P. and P.M. Allred. 1985. Replicability of Aquatic Multispecies Test Systems. In: Multispecies Toxicity Testing. J. Cairns, Jr. (ed.). SETAC Special Publication Series. Pergamon Press, Toronto.

Gilbert, R. and J.R. Glew. 1985. A portable percussion coring device for lacustrine and marine sediments. J. Sed. Petrol. 55: 607-608.

Gilbert, R.O. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold Co., New York.

Glew, J.R. 1988. A portable extruding device for close interval sectioning of unconsolidated core samples. J. Paleolim. 1: 235-239.

Glew, J.R. 1989. A new trigger mechanism for sediment samplers. J. Paleolim. 2: 241-243.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley and Sons, New York.

Grothe, D.R. and R.A. Kimerle. 1985. Inter- and intralaboratory variability in Daphnia magna effluent toxicity test results. Environmental Toxicology and Chemistry 4: 189-192.

Hart, D.R. and P.M. McKee. 1985. Study to Assess the Distribution of Radionuclides in the Aquatic Ecosystem in the Vicinity of the Serpent River Mouth. Report to Environment Canada, Ontario Region, by Beak Consultants Limited, Brampton, Ontario.

Helsel, D.R. 1986. Estimation of distributional parameters for censored water quality data. In: A.H. El-Shaarawi et al. (Eds.). Statistical Aspects of Water Quality Monitoring. Elsevier, Amsterdam.

Hensler, G.L. 1984. An Introduction to Statistical and Ecological Software, In: Gertz, S.M. and M.D. London (Eds.), Statistics in Environmental Sciences, STP 845, American Society for Testing and Materials, Philadelphia.

Hilsenhoff, W.L. 1981. Aquatic Insects of Wisconsin. Nat. Hist. Council Wisc., Madison. 60 p.

Hiltunen, J.K. and D.J. Klemm. 1980. A Guide to the Naididae of North America. U.S. Environmental Protection Agency Research Reports, Environmental Monitoring Series 600/4-80-031. Cincinnati, OH.

Hong, W., P.G. Meier and R.A. Deininger. 1988. Estimation of a single probit line from multiple toxicity test data. Aquatic Toxicology 12: 193-202.

Hunt, D.T. and A.L. Wilson. 1986. The Chemical Analysis of Water. General Principles and Techniques. Second Edition. Royal Society of Chemistry, London.

IJC. 1988. Quality Management Plan for Great Lakes International Surveillance with Implementation Guidelines. May 1988. International Joint Commission (draft).

ISO. 1982. Water Quality - Determination of the Inhibition of the Mobility of Daphnia magna Straus (Cladocera, Crustacea). International Standards Organization. ISO-6341-1982(E).

ISO. 1984a. Water Quality - Determination of the Acute Lethal Toxicity of Substances to a Freshwater Fish (Brachydanio rerio Hamilton-Buchanan (Teleostei, Cyprinidae)). Parts 1-3. International Standards Organization. ISO-7346-1984(E).

ISO. 1984b. Water Quality-Evaluation in an Aqueous Medium of the 'Ultimate' Aerobic Biodegradability of Organic Compounds - Method by Analysis of Dissolved Organic Carbon (DOC). International Standards Organization. ISO-7827-1984(E).

Jongman, R.H.G., C.J.F. ter Braak, and O.F.R. van Tongeren. 1987. Data analysis in community and landscape ecology. Pudoc, Wageningen, The Netherlands. 299 pp.

Kingston, J.C. 1986. Diatom Analysis. In: Charles and Whitehead, 1986b. pp. 6-1 to 6-11.

Kingston, J.C., J.P. Smol, S.S. Dixit and B.F. Cumming. 1988. Quality Assurance Plan - Paleoecological assessment of changes in Adirondack lake pH and alkalinity: Recent (since 1970) and historical (since 1850) trends. Queen's University, Kingston, Ont.

Klemm, D.J. (Ed.). 1985. Guide to the Freshwater Annelida (Polychaeta, Naidid, Tubificid, Oligochaeta and Hirudinea) of North America. Kendall/Hunt Publishing Co., Dubuque, Iowa.

Korndoerfer, L., Jr. 1984. The New Jersey Toxicity Test Laboratory Certification Program. New Jersey Administration Code (NJAC) 7:14-A1.1, October.

Kreis, R.G., Jr. 1986. Variability Study. In: Charles, D.F. and D.R. Whitehead (eds). Section 17. Paleoecological Investigation of Recent Lake Acidification (PIRLA): Methods and project description. Electric Power Research Institute, Palo Alto, CA.

Kreis, R.G., Jr. 1989. Variability Study - Interim results. Section 4 In: Charles, D.F. and D.R. Whitehead (eds.). Paleoecological Investigation of Recent Lake Acidification (PIRLA): 1983-1985: Interim Report, October 1989. Electric Power Research Institute, Palo Alto, CA.

Lanno, R.P. 1989. Nutritional considerations in toxicity testing. A short course presented at the 10th Annual SETAC Meeting, Toronto, Ontario.

LRTAP. 1986. Long-Range Transport of Air Pollutants. Quality Management Plan. Federal/Provincial Research and Monitoring Coordinating Committee.

Livingstone, D.A. 1955. A lightweight piston sampler for lake deposits. Ecology 36: 137-139.

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology: A Primer on Methods and Computing. John Wiley and Sons, New York.

MOE. 1983. Protocol to Determine the Acute Lethality of Liquid Effluents to Fish (G. Craig et al.). Toxicity Unit, Ontario Ministry of the Environment, Rexdale.

MOE. 1984. Principles of Control Charting. D.E. King and G.C. Ronan, Laboratory Services Branch, Data Quality Report Series. Ontario Ministry of the Environment, Rexdale.

MOE. 1985a. Laboratory Services Branch, Quality Management Plan. Ontario Ministry of the Environment.

MOE. 1985b. A Guide to the Collection and Submission of Samples for Laboratory Analysis. Ontario Ministry of the Environment, Laboratory Services Branch, Rexdale.

MOE. 1988a. Estimation of Analytical Detection Limits (MDL). Ontario Ministry of the Environment.

MOE. 1988b. Daphnia magna Acute Lethality Toxicity Test Protocol. (D.G. Poirier et al.). Aquatic Toxicity Unit, Ontario Ministry of the Environment, Rexdale.

MacGregor, D. and K. Doe. 1987. Basic elements of quality [illegible]ce in ecotoxicological testing. In: G.A. Sergy (Ed.). Recommendations [illegible]quatic Biological Tests and Procedures for Environmental Protection, C&r, DOE. Conservation and Protection, Environment Canada.

Mackereth, F.J.J. 1958. A portable core sampler for lake deposits. Limnol. Oceanogr. 3: 181-191.

Mackie, G.L., D.S. White and T.W. Zdeba. 1980. A Guide to Freshwater Mollusks of the Laurentian Great Lakes, with Special Emphasis on the Genus Pisidium. EPA-600/3-80/068. U.S. EPA. Duluth Environmental Research Laboratory, Duluth, Minnesota.

McAndrews, J.H., A.A. Berti and G. Norris. 1973. Key to the Quaternary pollen and spores of the Great Lakes region. Royal Ontario Museum. Misc Publ. 61 p.

McKean, C.J.P. and R.N. Nordin. 1986. A simple semi-co[illegible] piston corer for organic sediments. Hydrobiologia 137: 251-256.

McKee, P.M., W.J. Snodgrass, D.R. Hart, H.C. Duthie, J.H. McAndrews and W. Keller. 1987. Sedimentation rates and sediment core profiles of 238-U and 232-Th decay chain radionuclides in a lake affected by uranium mining and milling. Can. J. Fisheries and Aquatic Sci. 44: 390-398.

McLaughlin, D.L., W.I. Gizyn, D.E. Corrigan, W.D. McIlvean, R.G. Pearson and R. Aroup. 1988. A Numerical Decline Index Rating System to Monitor Changes in Tree Condition of Hardwood Forest Species. Proceedings, Technology Transfer Conference, 1988, Environment Ontario.

Merritt, R.W. and K.W. Cummins. 1984. An Introduction to the Aquatic Insects of North America (2nd Edition). Kendall/Hunt Publ. Co., Dubuque, Iowa. 721 p.

Mount, D.I. and L. Anderson-Carnahan. 1988. Methods for Aquatic Toxicity Identification Evaluations: Phase I, Toxicity Characterization Procedures. U.S. Environmental Protection Agency, Office of Research and Development. EPA-600-3-88-034, Duluth, MN.

Mount, D.I. and L. Anderson-Carnahan. 1988. Methods for Aquatic Toxicity Identification Evaluations: Phase II, Toxicity Identification Procedures. U.S. Environmental Protection Agency, Office of Research and Development. EPA-600-3-88-035, Duluth, MN.

Mount, D.I. and L. Anderson-Carnahan. 1988. Methods for Aquatic Toxicity Identification Evaluations: Phase III, Toxicity Confirmation Procedures. U.S. Environmental Protection Agency, Office of Research and Development. EPA-600-3-88-036, Duluth, MN.

Mueller-Dombois, D. and H. Ellenberg. 1974. Aims and Methods of Vegetation Ecology. John Wiley & Sons, New York.

Munro, M.A.R., A.M. Kreiser, R.W. Battarbee, S. Juggins, A.C. Stevenson, D.S. Anderson, N.J. Anderson, F. Berge, H.J.B. Birks, R.B. Davis, R.J. Flower, S.C. Fritz, E.Y. Haworth, V.J. Jones, J.C. Kingston, I. Renberg. 1990. Diatom quality control and data handling. Phil. Trans. R. Soc. Lond. B. (in press).

NCASI. 1990. Guidelines for the Selection of Laboratories to Perform Chronic Toxicity Tests Using Ceriodaphnia dubia and Fathead Minnows, (Pimephales promelas). Technical Bulletin No. 589. National Council for Air and Stream Improvement. New York.

NYDEC. 1985. Manual for Toxicity Testing of Industrial and Municipal Effluents. New York State Department of Environmental Conservation and Protection, New York.

OECD. 1981a. Decision of the Council Concerning Mutual Acceptance of Data in the Assessment of Chemicals. Annex 2. OECD Principles of Good Laboratory Practice. In: Guidelines for Testing of Chemicals. Organization for Economic Cooperation and Development.

OECD. 1981b. Guidelines for Testing of Chemicals. Organization for Economic Cooperation and Development, Paris.

OECD. 1989. Draft Council Decision-Recommendation on Compliance with Principles of Good Laboratory Practice (restricted). Organization for Economic Cooperation and Development.

Oliver, D.R. and M.E. Roussel. 1983. The genera of larval midges of Canada. In: The Insects and Arachnids of Canada, Part II. Agriculture Canada Publication 1746. Ottawa, Canada. 263 p.

Olsson, U. 1986. Radiometric dating. In: Berglund, B. (ed.), pp. 273-312.

PACE. 1989. Interlaboratory Comparison of the Daphnia magna Acute Lethality Bioassay. PACE Report No. 89-4. Petroleum Association for Conservation of the Canadian Environment.

Palmateer, G.A., B.J. Dutka, E.M. Janzen, S.M. Meissner and M. Sakellaris. 1989. Coliphages and bacteriophages in Canadian drinking waters. NWRI Contrib. #89-148. National Water Research Institute, Burlington.

Peck, D.V., J.L. Engels, K.M. Howe and J.E. Pollard. 1988. Aquatic Effects Research Program, Episodic Response Project, Integrated Quality Assurance Plan. EPA 600/x-88-274. U.S. EPA Environmental Monitoring Systems Laboratory, Las Vegas, Nevada.

Pennak, R.W. 1978. Freshwater Invertebrates of the United States. 2nd Ed. Wiley and Sons, Inc., NY.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross and R.M. Hughes. 1989. Rapid Bioassessment Protocols for use in Streams and Rivers: Benthic Macroinvertebrates and Fish. EPA/444/4-89-001. U.S. Environmental Protection Agency. Washington, D.C.

Prier, J.E., J.T. Bartola and H. Friedman. 1973. Quality Control in Microbiology. University Park Press. Baltimore.

QAMS. 1980. Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans. Quality Assurance Management Staff, Environmental Protection Agency, Washington. EPA-600/4-83-084. PB-83-170514.

QAMS. 1986. Development of Data Quality Objectives. Description of Stages I and II. Quality Assurance Management Staff, Environmental Protection Agency, Washington (draft).

Renberg, I. 1981. Improved methods for sampling, photographing and varve-counting of varved lake sediments. Boreas 10: 255-258.

Reteuna, C., P. Vasseur and R. Cabridenc. 1989. Performances of three bacterial assays in toxicity assessment. In: Environmental Bioassay Techniques and their Application, Munawar, M., G. Dixon, C.I. Mayfield, T. Reynoldson and M.H. Sadar (eds.). Hydrobiologia 188/189: XI-XII.

Saarnisto, M. 1986. Annually laminated lake sediments. In: Berglund, B. (ed.), pp. 343-370.

Santschi, P.H. 1985. The MERL mesocosm approach for studying sediment-water interactions and ecotoxicology. Env. Tech. Letters. 6: 335-350.

Scott, W.B. and E.J. Crossman. 1973. Freshwater Fishes of Canada. Bulletin 184. Fisheries Research Board of Canada, Ottawa.

Simola, H. 1986. Diatom responses to acidification and lime treatment in a clear-water lake: Comparison of two methods of analysis of a diatom stratigraphy. In: Smol, J.P., R.W. Battarbee, R.B. Davis, and J. Meril inen (eds.). Diatoms and Lake Acidity. Dr W. Junk Publishers, Dordrecht.

Smol, J.P. 1983. Paleophycology of a high arctic lake near Cape Herschel, Ellesmere Island. Can. J. Bot. 61: 2195-2204.

Smol, J.P. 1986. Chrysophytes. In: Charles and Whitehead (1986b), pp. 7-1 to 7-13.

Smol, J.P., B.F. Cumming, S.S. Dixit and J.C. Kingston. 1989. Coring equipment and procedures used for lakes in the Adirondacks -PIRLA II. Unpublished report no. 1., Queen's University, Kingston, Ontario.

Stephan, C.E. 1977. Methods for Calculating an LC50. Aquatic Toxicology and Hazard Evaluation. ASTM STP 634. (F.L. Mayer and J.L. Hanelink, Eds.). American Society for Testing and Materials, Philadelphia.

Stevenson, R.J. and R.L. Lowe. 1986. Sampling and Interpretation of Algal Patterns for Water Quality Assessments. In: Isom, B.G. (Ed.), Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems, STP 894. American Society for Testing and Materials, Philadelphia.

Stewart, K.W. and B.P. Stark. 1988. Nymphs of North American Stonefly Genera (Plecoptera). The Thomas Say Foundation Vol. XII. Entomological Society of America. 460 p.

Stimpson, K.S. and D.J. Klemm. 1982. A Guide to the Freshwater Tubificidae of North America. U.S. Environmental Protection Agency Research Reports, Environmental Monitoring Series 600/3-82-033. Cincinnati, OH.

Stockmarr, J. 1970. Tablets with spores used in absolute pollen analysis. Pollen et Spores 13: 615-621.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers, Inc., Chelsea, Michigan.

ter Braak, C.J.F. 1987. CANOCO - a FORTRAN program for canonical community ordination by partial detrended canonical correspondence analysis, principal components analysis and redundancy analysis (version 2.1) TNO Institute of Applied Computer Science, Wageningen, The Netherlands.

USDA. 1983. Techniques for Monitoring the Environmental Impact of Insecticides on Aquatic Ecosystems. Agricultural Handbook No. 613. U.S. Department of Agriculture, Forest Service.

Walker, E.M. and P.S. Corbet. 1975. The Odonata of Canada and Alaska. Vol. 3: Part III: The Anisoptera - Three Families. University of Toronto Press, Toronto, Ontario. 308 p.

Wiederholm, T. (Ed.). 1983. Chironomidae of the Holarctic region. Part I: Larvae. Ent. Scand. Suppl. 19: 1-457.

Wiggins, G.B. 1977. Larvae of the North American Caddisfly Genera (Trichoptera). University of Toronto Press. 401 p.

Wright, H.E., Jr. 1967. A square-rod piston sampler for lake sediments. J. Sed. Petrol. 37: 975-976.

Wright, H.E., Jr., E.J. Cushing, and D.A. Livingstone. 1965. Coring devices for lake sediments. In: B. Kummel and D. Raup (eds.). Handbook of Paleontological Techniques. Freeman and Company, San Francisco. pp. 494-520.

Youden, W.J. 1975. Statistical Techniques for Collaborative Tests. In: Statistical Manual of the AOAC. Association of Official Analytical Chemists, Arlington.

APPENDIX I

Specific Examples

TABLE A3.1: EXAMPLES OF SAMPLING DESIGNS
(see Gilbert, 1987, or Atkinson, 1985)

Simple Random Design: no stratification recognized

n random observations of variable X

mean $\bar{x}$ = $\frac{\sum X}{n}$

variance S^2 = $\frac{\sum (X - \bar{x})^2}{n-1}$

variance of mean $S^2_{\bar{x}}$ = $\frac{S^2}{\sqrt{n}}$

Stratified Random Design:

strata, each with size W_i, mean $\bar{x}_i$, variance S_i^2

mean $\bar{\bar{x}}$ = $\frac{\sum W_i \bar{x}_i}{\sum W_i}$

variance of mean $S^2_{\bar{x}}$ = $\frac{\sum W_i^2 S_i^2}{\sum W_i^2}$

TABLE A3.2: ESTIMATION OF REPLICATES REQUIRED
(see Gilbert, 1987, or Elliot, 1971)

Objective: Standard deviation of mean organism density at a station less than or equal to 20% of mean density:

$$\frac{S}{\sqrt{n}\ \bar{x}} \leq 0.2$$

Required number of replicates therefore is:

$$n \geq \frac{S^2}{\bar{x}^2 (.2)^2}$$

n = number of replicates
S = standard deviation (preliminary estimate)
$\bar{x}$ = station mean (preliminary estimate)

TABLE A3.3: A DQO STATEMENT FORMAT FOR A COMPLIANCE STUDY

The Domain:

Samples collected along the edge of a specified mixing zone over a period of one month, the allowed averaging period.

The Summary Statistic:

A mean concentration of $\bar{C}_{MZ}$ representing the edge of the mixing zone over the defined period.

The Decision:

Does $\bar{C}_{MZ}$ significantly exceed the water quality objective C_{OB}?

The Data Quality Objective (DQO):

- To be able to identify a mean concentration exceeding the water quality objective by ΔC as a significant exceedance (i.e., to achieve a resolution of ΔC).
- The confidence interval (imprecision) of $\bar{C}_{MZ}$ is: $t_\alpha S / \sqrt{n}$,
  where: t is the critical value of Student's t with a Type I error probability of α,
  S is the standard deviation of concentration based on analytical, sampling method and environmental variance ($S^2 = S_a^2 + S_s^2 + S_e^2$), and
  n is the number of samples contributing to the mean.
- The upward bias in $\bar{C}_{MZ}$ is assumed to be: b.
- The cost constraint on n is: $n < \text{Budget}/(D_a + D_s)$,
  where: D_a is the analytical cost per sample, and
  D_s is the sampling cost per sample.
- The DQO requires that: $(t_\alpha S / \sqrt{n}) + b < \Delta C$. For stated values of α, n and ΔC, achievable limits of S and b consistent with this equation are established. Based on expected environmental variance S_e, S may be partitioned to set additional limits for S_a and S_s.
- Based on the water quality objective, a minimum reporting level is defined (e.g., $C_{OB}/10$) above which all samples should be quantitative.

TABLE A3.4: A DQO STATEMENT FORMAT FOR A SPATIAL TREND STUDY

The Domain:

Surface samples collected in August from open water areas of Lakes 1 and 2 are compared.

The Summary Statistic:

The mean concentrations $\bar{C}_1$, and $\bar{C}_2$ representing the two lakes over the same time period, and their difference $\bar{C}_1 - \bar{C}_2$.

The Decision:

Do $\bar{C}_1$ and $\bar{C}_2$ differ significantly?

The Data Quality Objective (DQO):

- To be able to identify a difference of ΔC between means as significant (i.e., to achieve a resolution of ΔC).
- The confidence interval (imprecision) of $\bar{C}_1 - \bar{C}_2$ is: $t_\alpha \, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

  where: t is the critical value of Student's t with a Type I error probability of α,

  S_p is the pooled standard deviation of concentration in the two lakes ($S_p^2 = (n_1-1)S_1^2+(n_2-1)S_2^2 \; / \; (n_1+n_2-2)$ based on analytical, sampling and environmental variance ($S_1^2 = S_{1a}^2 + S_{1s}^2 + S_{1e}^2$; $S_2^2 = S_{2a}^2 + S_{2s}^2 + S_{2e}^2$), and

  n_1 and n_2 are the number of samples contributing to each lake mean.
- The upward bias in $\bar{C}_1$ relative to that in $\bar{C}_2$ is: b.
- The cost constraint on $n = n_1 + n_2$ is: $n < \text{Budget}/(D_a + D_s)$,

  where: D_a is the analytical cost per sample, and

  D_s is the sampling cost per sample.
- The DQO requires that: $\left(t_\alpha \, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) + b < \Delta C$.

  For stated values of α, n_1, n_2 and ΔC, achievable limits of S_p and b consistent with this equation are established. Based on expected environmental variance S_{pe}, S_p may be partitioned to set additional limits for S_{pa} and S_{ps}.
- Based on expected concentrations over the domain of study, a minimum reporting level is defined, above which all samples should be quantitative.

TABLE A4.1a: RECOMMENDED PREPARATION OF RECONSTITUTED FRESH WATER (EPA, 1979b)

	Salts Required (mg/L)						
Name	$NaHCO_3$	$CaSO_4.2H_2O$	$MgSO_4$	KCl	pH[1]	Hardness[2]	Alkalinity[2]
Very soft	12	7.5	7.5	0.5	6.4-6.8	10-13	10-13
Soft	48	30.0	30.0	2.0	7.2-7.6	40-48	30-35
Hard	192	120.0	120.0	8.0	7.6-8.0	160-180	110-120
Very hard	384	240.0	240.0	16.0	8.0-8.4	280-320	225-245

[1] Approximate equilibrium pH after aeration and with fish in water.
[2] Expressed in mg/L as $CaCO_3$.

TABLE A4.1b: SPECIFICATIONS FOR ALTERNATIVE DILUTION WATER (EPA, 1978)

Contaminant	Maximum Concentration
Suspended Solids	20 mg/L
Total Organic Carbon	10 mg/L
Un-ionized Ammonia	20 ug/L
Residual Chlorine	3 ug/L
Total Organophosphorus Pesticides	50 ng/L
Total Organochlorine Pesticides (including PCBs)	50 ng/L

TABLE A4.1c: RECOMMENDED PREPARATION OF RECONSTITUTED SEA WATER (EPA, 1979b)

Chemical	Amount[1]	Chemical	Amount[1]	Chemical	Amount[1]
NaF	3 mg	KCl	700 mg	NaCl	23.50 g
$SrCl_2.6H_2O$	20 mg	$CaCl_2 \cdot 2H_2O$	1.47 g	$Na_2SiO_3 \cdot 9H_2O$	20 mg
H_3BO_3	30 mg	Na_2SO_4	4.00 g	Na_4EDTA	1 mg
KBr	100 mg	$MgCl_2 \cdot 6H_2O$	10.78 g	$NaHCO_3$	200 mg

[1] Amount of reagent grade chemical added to 890 mL distilled water.

TABLE A5.1: AUTOCLAVE STERILIZATION TEMPERATURE AND TIME
(after Bordner, 1985; APHA, 1987)

Material	Temperature and Time
Membrane filters and pads	121°C for 10 min (*)
Carbohydrate-containing media (lauryl tryptose, BGB broth, etc.)	121°C for 12 to 15 min not to exceed 45 min in autoclave remove as soon as possible
Contaminated materials and discarded cultures	121°C for 30 min
Membrane filter assemblies (wrapped), sample collection bottles (empty)	121°C for 15 min and dried
Dilution water (99 mL in screw-cap bottles)	121°C for 15 min
Rinse water volumes of 500 to 1,000 mL	121°C for 30 min
Rinse water in excess of 1 L	121°C with time adjusted for volume; check for sterility

* 121°C is the temperature of steam at 15 PSI; time is selected to ensure three to four decimal reductions in spore count.

TABLE A5.2: SPECIFICATIONS FOR MEMBRANE FILTERS AND ABSORBENT PADS (after Bordner, 1985)

Property	Specification
Membrane Filter (0.45 u)	
Diffusibility	Laboratory-pure water should diffuse uniformly through the filters in 15 s, and the filters should have no dry spots when floated on water.
Flow rate	Flow rate should be at least 55 mL/min/cm^2 at 25°C and a differential pressure of 70 cm mercury.
Retention	Filters should retain a 100 mL suspension of Serratia marcescens containing $1x10^3$ cells/mL.
Inhibition	Filters should not inhibit growth of the test organism. When five aliquots of the same sample are analyzed by the MF and spread plate method using the same agar medium, the arithmetic mean of the MF counts should be at least 90% of the arithmetic mean of the plate count.
Extractables	Water extractables from an MF should not exceed 2.5% after boiling in 100 mL laboratory-pure water (Table A5.5) for 20 min, drying, cooling and bringing to constant weight.
Toxicity	Filters should be non-toxic and free of bacterial growth-inhibiting or growth-stimulating substances.
Sterility*	Growth should not result when an MF is placed on a pad saturated with tryptone glucose extract (TGE) broth, nutrient broth, or TGE agar, and incubated at 35°C for 24 h.
Absorbent Pads	
Absorptivity	2.0 ± 0.2 mL broth
Acidity	Pads should release less than 1 mg total acidity calculated as $CaCO_3$ when titrated to the phenolphthalein endpoint with 0.02N NaOH.

* Samples passing this test could be contaminated with Legionella, Clostridium, or viable non-replicable organisms.

TABLE A5.3: MONITORING LABORATORY EQUIPMENT (after Bordner, 1985)

Equipment	Monitoring Procedure
Incubator (air/water-jacket)	o If partially submersible glass thermometer is used, bulb and stem must be immersed in liquid/water to the mark on stem. o Measure and record temperature twice daily on top and bottom shelves. Periodically measure temperature on all shelves in use. o Expand monitoring locations proportionately for walk-in incubators. o Recording thermometer and alarm system are recommended.
Water Bath	o Check and record temperature daily. Bath must maintain the uniform temperature needed for the test in use. o Maintain accurate thermometer completely immersed in water bath, and record temperature daily. o A recording thermometer (graduated to 0.1°C) and alarm system are recommended. o Clean monthly.
Autoclave	o Record temperature and pressure for each run. Recording thermometer recommended. o Verify that autoclave maintains uniform operating temperature. o Test performance with spore strips or suspensions each run. If evidence of contamination occurs, check until the cause is identified and eliminated. o Procure semi-annual preventive maintenance inspections.
Hot Air Oven	o Test performance with spore strips or suspensions quarterly. o Equip and monitor sterilization with a thermometer accurate in the 160 to 180°C range, and record temperature daily.
Refrigerator at 4°C	o Check and record temperature daily. o Clean monthly. o Require identification and dating of all material. Defrost unit and discard outdated materials in refrigerator and freezer compartments every three months.
Freezers	o Check temperature and record daily. o Use of recording thermometer and alarm system recommended. o Require identification and dating of all materials o Clean and defrost freezer every six months. Discard outdated materials.

TABLE A5.3: MONITORING LABORATORY EQUIPMENT (after Bordner, 1985)

Equipment	Monitoring Procedure
Balance	o An analytical balance with a sensitivity of 1 mg or less at a 10-g load should be used for weighing 2 g or less. For larger quantities, a balance with accuracy of 50 mg at a 150-g load should be used. o Check balance monthly with a set of certified Class S weights or weights of equivalent accuracy. o Wipe balance and weights clean after each use. o Protect weights from laboratory atmosphere and corrosion. o Contract with a qualified expert for balance maintenance on an annual basis.
pH Meter	o Compensate for temperature with each use. o Date standard buffer solutions when first opened and check monthly with another pH meter. Discard the buffer solution if the pH is more than ± 0.1 pH unit from the manufacturer's stated value or if it is contaminated with microorganisms. o Standardize with at least two standard buffers (pH 4.0, pH 7.0, or pH 10.0) before each use. o Do not re-use buffer solution. o Have meters inspected at least yearly as part of maintenance contract.
Water Deionizer	o Monitor water for conductance daily. Monitor trace metals and other toxic or nutritive compounds annually. o Replace cartridges as indicated by manufacturer or as indicated by analytical results. o Monitor bacterial counts at exit point of unit. Replace the cartridges when standard plate count exceeds 100/mL.
Water Still	o Drain and clean monthly according to instructions from the manufacturer. o Drain and clean distilled water reservoir quarterly. o Check distilled water continuously or daily using a conductance meter. o Conduct chemical and biological tests on water to detect toxicity or stimulation effect annually. o Conduct standard plate counts monthly on stored water and clean out reservoir if count greater than 100/mL.
Washing Machine	o Drain and clean reservoir weekly. o Conduct chemical test of rinse water monthly. o Conduct inhibitory residual test annually.

TABLE A5.4: MAXIMUM STORAGE TIMES FOR PREPARED MEDIA
(after Bordner, 1985; APHA, 1987)

Medium	Holding Time Limit
Before Preparation	
Sealed Container	2 years
After Opening	6 months
After Preparation	
Membrane filter (MF) broth in screw-cap flasks at 4°C	96 hours
MF agar in plates with tight-fitting covers at 4°C	2 weeks
Agar or broth in loose-cap tubes at 4°C (if kept in humidified refrigerator)	1 week (1 month)
Agar or broth in tightly closed screw-cap tubes at 4°C	3 months
Poured agar plates with loose-fitting cover in sealed plastic bags at 4°C	2 weeks
Large volume of agar in tightly closed screw-cap flask or bottle at 4°C	3 months

TABLE A5.5: CRITERIA FOR LABORATORY PURE WATER USED IN MICROBIOLOGICAL PROCEDURES (after Bordner, 1985; APHA, 1987)

Test	Monitoring Frequency	Limit
Routine Tests		
Conductivity	continuously or with each use	< 0.5 M resistance or < 2 S/cm at 25°C
pH	with each use	5.5 to 7.5
Standard plate count	monthly	< 100 bacteria/mL
Water use test	monthly and for a new source	no significant change in plate count (≥ 5 plates)
Follow-up Tests When Problems Are Indicated		
Total organic carbon		< 0.5 mg/L
Heavy metals, individually (Cd, Cr, Cu, Ni, Pb, Zn)		< 0.05 mg/L
Heavy metals, total (Cd, Cr, Cu, Ni, Pb, Zn)		≤ 0.1 mg/L
Ammonia/organic nitrogen		< 0.01 mg/L
Total chlorine residual		< 0.5 mg/L
Water suitability test * (inhibition or promotion of E. aerogenes)		test/control ratio: 0.8 to 3.0

* Water suitability test is highly sensitive; failure does not necessarily indicate impact on routine laboratory analyses.

TABLE A5.6: CONTROL CULTURES FOR STANDARD MICROBIOLOGICAL TESTS (after Bordner, 1985)

Parameter	Positive	Negative
Total coliforms	Escherichia coli Enterobacter aerogenes	Staphylococcus aureus Pseudomonas sp.
Fecal coliforms	E. coli	E. aerogenes Streptococcus faecalis
Fecal streptococci	Streptococcus faecalis	Staphylococcus aureus E. coli Streptococcus mutans

APPENDIX 2

Quality Assurance Glossary

DRAFT

QUALITY ASSURANCE GLOSSARY AND ACRONYMS

Quality Assurance Management Staff (RD-680)
Office of Modeling, Monitoring Systems and Quality Assurance
Office of Research and Development
United States Environmental Protection Agency
401 M Street, SW
Washington, DC 20460

July 23, 1990

DRAFT

GLOSSARY OF TERMS

Absolute method - a body of procedures and techniques for which measurement is based entirely on physically defined, fundamental quantities.

Acceptable quality level - a limit above which quality is considered satisfactory and below which it is not. In sampling inspection, the maximum percentage of defects or failures that can be considered satisfactory as an average.

Acceptable quality range - the interval, between specified upper and lower limits of a sequence of values, within which the values are considered to be satisfactory.

Acceptable value - an observed or corrected value that falls within the acceptable range. See Corrected value and Observed value.

Acceptance sampling - the procedure of drawing samples from a lot or population to determine whether to accept or reject a sampled lot or population.

Accepted reference value - a numerical quantity that serves as an agreed-upon basis for comparison, and which is derived as; 1) a theoretical or established quantity based on scientific principles, 2) an assigned value, based on experimental work of some recognized organization, or 3) a consensus quantity based on collaborative experimental work under the auspices of a scientific or engineering group.

Accreditation - a formal recognition that a laboratory is competent to carry out specific tests or specific types of tests. See Certification.

Accreditation criterion - a requirement that a laboratory must meet to receive authorization and approval to perform a specified task.

Accredited laboratory - a laboratory which has been evaluated and given approval to perform a specified measurement or task, usually for a specific parameter and a specified period of time.

Accuracy - the degree of agreement between an observed value and an accepted reference value. Accuracy includes a combination of random and systematic error or bias components which are due to sampling and analytical operations; a data quality indicator.

Aliquant - a subsample derived by a divisor that divides a sample into a number of equal parts but leaves a remainder; a subsample resulting from such a divisor. See Subsample.

Aliquot - a subsample derived by a divisor that divides a sample into a number of equal parts and leaves no remainder; a subsample resulting from such a division. In analytical chemistry the term aliquot is generally used to define any representative portion of the sample.

Alpha error - see "Type I Error."

DRAFT

Batch-lot - the samples collected under sufficiently uniform conditions to be processed as a group. See Batch, Batch size.

Batch-sample - one of the samples drawn from a batch.

Batch-size - the number of samples in a batch-lot.

Beta error - see Type II Error.

Bias - systematic error manifested as a consistent positive or negative deviation from the known or true value. It differs from random error which shows no such deviation.

Blank sample - a clean sample or a sample of matrix processed so as to measure artifacts in the measurement (sampling and analysis) process.

Blind sample - a subsample submitted for analysis with a composition and identity known to the submitter but unknown to the analyst and used to test the analyst's or laboratory's proficiency in the execution of the measurement process. See Double-blind sample.

Bulk sample - a sample taken from a larger quantity (lot) for analysis or recording purposes.

Calibrant - see Calibration standard.

Calibrate - to determine, by measurement or comparison with a standard, the correct value of each scale reading on a meter or other device, or the correct value for each setting of a control knob. The levels of the calibration standards should bracket the range of planned measurements.

Calibration-check - see Calibrate.

Calibration-check standard - see Calibration standard.

Calibration curve - the graphical relationship between the known values for a series of calibration standards and instrument responses.

Calibration drift - the difference between the instrument response and a reference value after a period of operation without recalibration.

Calibration standard - a substance or reference material used to calibrate and instrument.

Candidate method - a body of procedures and techniques of sample collection and/or analysis that is submitted for approval as a reference method, an equivalent method, or an alternative method.

Carrying-agent - any diluent or matrix used to entrain, dilute or to act as a vehicle for a compound of interest.

DRAFT

Collocated sample - one of two or more independent samples collected so that each is equally representative for a given variable at a given space and time.

Collocated samplers - two or more identical instruments usually operated simultaneously, to supply a series of duplicate or replicate samples for estimating precision.

Comparability - a measure of the degree to which different methods, data sets and/or decisions can be represented as similar; a data quality indicator.

Completeness - the amount of valid data obtained compared to the planned amount, and usually expressed as a percentage; a data quality indicator.

Component of variance - a part of the total variance associated with a specified source of variation.

Composite sample - a sample prepared by physically combining two or more samples or substances and processed to ensure homogeneity. See Flow-proportioned sample and Time-proportioned sample.

Confidence interval - a range that has a designated probability (i.e., confidence level) of including a specific population parameter.

Confidence level - the probability, usually expressed as a percentage, that a confidence interval will include a specific population parameter; confidence levels usually range from 90 to 99 percent.

Control chart - a graph of some measurement plotted over time or sequence of sampling, together with control limit(s) and, usually, a central line. See Action limit, Central line, Control limit and Warning limit.

Control limit - a specified boundary on a control chart that, if exceeded, indicates a process is out of statistical control, and the process must be stopped, and corrective action taken before proceeding; the boundary specified is usually three standard deviations (i.e., the 99.72% confidence level.)

Control sample - see Quality control sample and Check sample.

Control standard - see Calibration-check standard.

Corrected value - the observed value after correction for values of a blank.

Correlation - a measure of association between two variables.
See Correlation coefficient.

Correlation coefficient - a number between -1 and 1 that indicates the degree of linearity between two sets of numbers. The closer to -1 or +1, the stronger the linear relationship between the sets (i.e., the better the correlation.) Values close to zero suggest no correlation between the sets and imply independence.

DRAFT

Discharge - any pollutant or combination of pollutants added to the environment from any point or area source.

Document control - a systematic procedure for indexing documents by number, date and revision number.

Double-blind sample - a sample submitted to evaluate performance with concentration and identity unknown to the analyst. See Blind sample.

Duplicate - an adjective describing the taking of a second sample or performance of a second measurement or determination. Often incorrectly used as a noun and substituted for "duplicate sample." Replicate is to be used if there are more than two items. See Replicate.

Duplicate analysis or measurement - the analysis or measurement of the variable of interest performed as identically as possible on two subsamples of a sample. Duplicate analysis is used to evaluate analytical or measurement variance.

Duplicate sample - one of two samples taken from the same population and carried through all steps of the sampling and analytical procedures in an identical manner. Duplicate samples are used to assess variance of the total method including sampling and analysis. See Split sample.

Dynamic blank - a sample-collection material or device (e.g., filter or reagent solution) that is not exposed to the material to be selectively captured but is transported and processed in the same manner as the sample. See Field blank, Instrumental blank and Sampling equipment blank.

Dynamic calibration - standardization of both the measurement and collection systems using a reference material similar to the unknown. For example, a series of air-mixture standards containing sulfur dioxide of known concentrations could be used to calibrate a sulfur dioxide bubbler system.

Environmental sample - a sample of any material that is collected from an environmental source.

Environmentally related measurement - any assessment of environmental concern generated through or for field, laboratory, or modelling processes; the value obtained from such an assessment.

Equivalent method - any method of sampling and/or analysis demonstrated to have a consistent and quantitatively known relationship to a reference method under specified conditions.

Error - any discrepancy between a computed, observed, or measured quantity and the true, specified, or theoretically correct value of that quantity.

Error (measurement) - the difference between an observed or corrected value of a variable and a specified, theoretically correct, or true value.

Error function - see Normal distribution.

Experimental variable - see Independent variable.

DRAFT

In-control - a condition indicating that performance of the quality control system is within the specified control limits, i.e., that a stable system of chance is operating and resulting in statistical control. See Control chart.

Independent variable - a variable that is set at pre-selected level when a controlled experiment is conducted.

Inspection criterion - the specification(s) and rationale for rejecting and accepting samples in a particular sampling plan.

Instrument blank - a clean sample processed through the instrumental steps of the measurement process; used to determine the instrument component of the measurement error. See Dynamic blank.

Interference - a positive or negative effect on a measurement caused by a variable other than the one being investigated.

Interference equivalent - the mass or concentration of a foreign substance which gives the same measurement response as one unit of mass or concentration of the substance being measured.

Interlaboratory calibration - the process, procedures, and activities for standardizing a given measurement system to ensure that laboratories participating in the same program can produce comparable data.

Interlaboratory precision - a measure of the variation among laboratories participating in the same test; usually the standard deviation of the data sets.

Interlaboratory test - a test performed by two or more laboratories on the same material for the purpose of assessing the capabilities of an analytical method or for comparing different methods.

Internal quality control - see Intralaboratory quality control.

Internal standard - a standard added to samples in known amount and carried through the procedure as a reference for calibration and controlling instrumental and analytical precision and bias.

Intralaboratory quality control - the routine activities and checks, such as periodic calibrations, duplicate analyses and spiked samples, included in normal internal procedures to control the accuracy and precision of measurements.

Intralaboratory precision - a measure of the method/sample specific analytical variation within a laboratory; usually the standard deviation of a data set. See also Standard deviation and Variance.

Laboratory accreditation - see Accredited laboratory and Accreditation.

Laboratory blank - see Reagent blank.

DRAFT

Matrix - a type of medium (e.g., surface water, drinking water) in which the analyte of interest may be contained. See Medium.

Matrix spike -see Spiked sample.

Matrix spike duplicate sample analysis - see Matrix, Duplicate analysis and Spiked sample.

Maximum contaminant level - the highest permissible concentration of a pollutant that may be delivered to any receptor.

Maximum holding time - the duration for which a sample can be kept under specified conditions without significant degradation of the analyte(s) or property of interest.

Mean - see Arithmetic mean.

Measured value - see Corrected value.

Measurement range - the range over which the precision and/or recovery of a measurement method are regarded as acceptable. See Acceptable quality range.

Measure of central tendency - a statistic that describes the grouping of values in a data set around some common value (e.g., the median, arithmetic mean, or geometric mean.)

Measure of dispersion - a statistic that describes the variation of values in a data set around some common value. See Range, Variance and Standard deviation.

Medium - a substance (e.g., air, water, soil) which serves as a carrier of the analytes of interest. See Matrix.

Medium blank - see Field blank and/or Laboratory blank.

Median - the middle value for an ordered set of n values; represented by the central value when n is odd or by the mean of the two most central values when n is even.

Method - a body of procedures and techniques for performing a task (e.g., sampling, characterization, quantification) systematically presented in the order in which they are to be executed.

Method blank - a clean sample processed simultaneously with and under the same conditions as samples containing an analyte of interest through all steps of the analytical procedure.

Method check sample - see Spiked laboratory blank.

Method detection limit (MDL) - in a given matrix and with a specific method, the minimum concentration of an analyte that can be identified, qualitatively or quantitatively measured, and reported to be greater than zero at the 99% confidence level.

Method of least squares - see Least squares method.

DRAFT

Primary standard - a substance or device, with a property or value that is unquestionably accepted (within specified limits) in establishing the value of the same or related property of another substance or device.

Probability - a numerical estimate of the chance that a particular event will occur. Its value can range from 0 (impossible) up to 1 (certain).

Procedure - a set of systematic instructions for performing an operation.

Proficiency testing - a systematic program in which one or more standardized samples is analyzed by one or more laboratories to determine the capability of each participant.

Property - a quality or trait belonging and peculiar to a thing; a response variable is a measure of a property. Synonym: Characteristic.

Protocol - a detailed written procedure for a field and/or laboratory operation (e.g., sampling, analysis.)

Quality - the sum of features and properties/characteristics of a product or service that bear on its ability to satisfy stated needs.

Quality assurance (QA) - an integrated system of activities involving quality planning, quality control, quality assessment, quality reporting and quality improvement to ensure that a product or service meets defined standards of quality with a stated level of confidence.

Quality Assurance Narrative Statement - a description of the quality assurance and quality control activities to be followed for a research project.

Quality Assurance Objectives - the upper and lower limiting values of the data quality indicators as defined by the data user's acceptable error bounds.

Quality Assurance Program Plan (QAPP) - a formal document describing the management policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation plan of an agency, organization or laboratory for ensuring quality in its products and utility to its users.

Quality Assurance Project Plan (QAPjP) - a formal document describing the detailed quality control procedures by which the data quality requirements defined for the data and decisions in a specific project are to be achieved.

Quality Circle - a small group of individuals from an organization or unit who have related interests and meet regularly to consider problems or other matters related to the quality of the product or process.

Quality control (QC) - the overall system of technical activities whose purpose is to measure and control the quality of a product or service so that it meets the needs of users. The aim is to provide quality that is satisfactory, adequate, dependable, and economical.

DRAFT

Records system (or plan) - a written, documented group of procedures describing required records, steps for producing them, storage conditions, retention period and circumstances for their destruction or other disposition.

Recovery efficiency - in an analytical method, the fraction or percentage of a target analyte extracted from a sample containing a known amount of the analyte.

Reference material - a material or substance, one or more properties of which are sufficiently well established to be used for the calibration of an apparatus, the assessment of a measurement method, or assigning values to materials.

Reference method - a sampling and/or measurement method which has been officially specified by an organization as meeting its data quality requirements.

Reference standard - see Calibration standard.

Relative standard deviation - see Coefficient of variation.

Reliability - the likelihood that an instrument or device will function under defined conditions for a specified period of time.

Repeatability - the capability of a method, test or instrumental procedure to replicate a result under specified conditions (e.g., by a single analyst and/or instrument over a specified period of time.)

Replicability - see Repeatability.

Replicate - an adjective referring to the taking of more than one sample or performance of more than one analysis. Incorrectly used as a noun in place of replicate analysis. Replicate is to be used when referring to more than two items. Duplicate is to be used when there are only two. See Duplicate.

Replicate analysis or measurement - the analysis or measurement of the variable of interest performed as identically as possible on two or more subsamples of a sample. Replicate analyses are used to assess analytical or measurement variance. Often incorrectly called "replicate sample" or "replicate". Replicate analysis is to be used when there are more than two and duplicate analysis if there are only two.

Replicate sample - one of two or more samples taken from the same population and carried through all steps of the sampling and measurement process in an identical manner. Replicate samples are used to assess total (sampling and analysis) method variance. Often incorrectly used in place of replicate analysis.

Representative sample - a sample taken so as to reflect the frequency distribution of the variable of interest in the population as accurately and precisely as specified. To ensure representativeness, the sample may be either completely random or stratified depending upon the conceptualized population and the objective (i.e., upon the decision to be made.)

DRAFT

Selectivity (analytical chemistry) - the capability of a method or instrument to respond to a target substance or constituent in the presence of nontarget substances.

Sensitivity - capability of method or instrument to discriminate between measurement responses of a variable of interest.

Significant digit - any of the digits 0 through 9, excepting leading zeros and some trailing zeros, which is used with its place value to denote a numerical quantity to a desired rounded number. See Rounded number.

Significant figure - see Significant digit.

Single operator precision - the standard deviation of the results of a series of determinations by the same analyst or operator, all other conditions being equal.

Site - the area within a boundaries established for a defined activity.

Span-drift - the change in the output of a continuous monitoring instrument over a stated time period during which the instrument is not recalibrated.

Span-gas - a gas of known concentration which is used routinely to calibrate the output level of an analyzer. See Calibration check standard.

Specimen - see Sample.

Spike - a known mass of target analyte added to a blank sample or subsample; used to determine recovery efficiency or for other quality control purposes.

Spiked laboratory blank - see Spiked reagent blank.

Spiked reagent blank - a reagent blank fortified with a known mass of the target analyte; usually used to determine the recovery efficiency of the method.

Spiked sample - a sample prepared by adding target or nontarget material to a matrix sample; used, for example, in an analytical method to determine the effect of the matrix on the recovery efficiency.

Spiked sample duplicate analysis - see Duplicate analysis and Spiked sample.

Split sample - one of two or more portions taken from a sample or subsample used to replicate the measurement of the variable(s) of interest.

Standard (measurement) - a substance or material with a property quantified with sufficient accuracy to permit its use to evaluate the same property in a similar substance or material. Standards are generally prepared by placing a reference material in a matrix. See Reference material.

DRAFT

Surrogate analyte - a pure substance with properties that mimic the analyte of interest. It is unlikely to be found in environmental samples and is added to them for quality control purposes.

Surveillance - the act of maintaining supervision of or vigilance over a well-specified portion of the environment so that detailed information is provided concerning the state of that portion.

Synthetic sample - see Quality control sample.

Systematic error - a consistent deviation in the results of sampling and/or analytical processes from the expected or known value. Such error is caused by human and methodological bias.

Systems error - See Total systems error.

Target - the chosen object of investigation for which qualitative and/or quantitative data or information is desired, e.g., the analyte of interest.

Technical systems audit - a thorough, systematic on-site, qualitative review of facilities, equipment, personnel, training, procedures, recordkeeping, data validation, data management, and reporting aspects of a total measurement system.

Technique - a principle and/or the procedure of its application for performing an operation.

Test - a procedure used to identify or characterize a substance or constituent. See Method.

Test data - see Data.

Test determination - the process of obtaining an observed value of a property for a single sample. Also, the observed value obtained (i.e., the test result.) See Test result.

Test method - see Method.

Test portion - a subsample of the proper amount for analysis and measurement of the property of interest. A test portion may be taken from the bulk sample directly, but often preliminary operations, such as mixing or further reduction in particle size, are necessary. See Subsample.

Test result - a product obtained from performing a test determination. See Test determination.

Test sample - see Test portion.

Test specimen - see Test portion.

Test unit - see Test portion.

DRAFT

Validated method - a method which has been determined to meet certain performance criteria for sampling and/or measurement operations.

Validation - the process of substantiating specified performance criteria.

Value - the magnitude of a quantity. A single piece of factual information obtained by observation or measurement and used as a basis of calculation.

Variable - an entity subject to variation or change.

Variance (statistical definition) - a measure of the dispersion of a set. The sum of the squares of the difference between the individual values of a set and the arithmetic mean of the set, divided by one less than the number of values in the set. (The square of the standard deviation.) See also Measure of dispersion.

Verifiable - the ability to be proven or substantiated.

Warning limit - a specified boundary on a control chart that indicates a process may be going out of statistical control and that certain precautions are required. The boundary specified is usually two standard deviations (i.e.,the 95% confidence level.)

Within-laboratory standard deviation - see Intralaboratory precision.

Working standard - see Secondary standard.

Zero drift - the change in instrument output over a stated time period of nonrecalibrated, continuous operation, when the initial input concentration is zero; usually expressed as a percentage of the full scale response.